

Portland State University PDXScholar

Dissertations and Theses

Dissertations and Theses

Spring 5-22-2018

Exploring District Judges' Decision Making in the Context of Admitting Expert Testimony

Andrew Bryan Dzeguze
Portland State University

Let us know how access to this document benefits you.

Follow this and additional works at: https://pdxscholar.library.pdx.edu/open_access_etds

 Part of the [Law Commons](#), [Legal Studies Commons](#), and the [Political Science Commons](#)

Recommended Citation

Dzeguze, Andrew Bryan, "Exploring District Judges' Decision Making in the Context of Admitting Expert Testimony" (2018).
Dissertations and Theses. Paper 4392.

[10.15760/etd.6276](https://pdxscholar.library.pdx.edu/open_access_etds/10.15760/etd.6276)

This Dissertation is brought to you for free and open access. It has been accepted for inclusion in Dissertations and Theses by an authorized administrator of PDXScholar. For more information, please contact pdxscholar@pdx.edu.

Exploring District Judges' Decision Making in the Context of Admitting
Expert Testimony

by

Andrew Bryan Dzeguze

A dissertation submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy
in
Public Affairs and Policy

Dissertation Committee:
Christopher Shortell, Chair
Phillip J. Cooper
Birol Yesilada
Robert L. Fountain

Portland State University
2018

© 2018 Andrew Bryan Dzeguze

Abstract

Over the last several decades, multiple schools of thought have emerged regarding what impacts judicial decision making. In contrast to the classic legal model, studies have argued alternatively that judges are policy actors who rule consistent with their political attitudes; that behavioral traits such as race, gender and socialization influence judicial conduct, both consciously and unconsciously; that whatever policy interests judges may have, these are moderated by institutional constraints and strategic considerations; and that judges are subject to some common cognitive shortcuts in decision making, although they may be moderated or present differently than in the general population in light of their training and experience.

Most of these studies, particularly in political science, have tended to focus on Supreme Court or appellate decisions on politically salient subject matter such as the scope of the Fourth Amendment or racial discrimination. The cognitive studies, by comparison, have primarily used experimental conduct, often with artificially extreme variations between legal and factual issues to assess the impact of legal training. Other than field review articles, most have focused on a single potential explanatory variable such as ideology, gender or legal training. To date, there has been very limited study of the more routine tasks judges engage in at the trial court level such as pre-trial evidentiary rulings or comparative assessments of the relative explanatory power of factors drawn from multiple approaches to decision making.

The present study involved both a qualitative and quantitative assessment of Federal district court decisions on the admissibility of expert witnesses. Employing thematic analysis of all cases involving a substantive analysis of this issue from 2010-

2015 in nine district courts, a default pattern emerged that judges are reluctant to exclude experts except in extreme cases. Moreover, judges appear to have adopted several practices consistent with minimizing the cognitive burden of decision making. These findings suggest that judges are acting consistently with legal norms and the broad outlines of legal precedent, but in a manner which may lead to sub-optimal outcomes in some circumstances. Quantitative analysis of the same data suggests that judges are subject to a variety of significant influences including legal precepts, political ideology and cognitive heuristics in different settings. Moreover, the influence of issues such as ideology appear to be associated with some courts and not others, with circuit level precedent being the most obvious intervening factor to explain the difference.

The circuit level impacts on behavior and several other findings in this study suggest that much more nuance is present than is normally acknowledged in the study of judicial decision making. The results of this study also suggest policy makers should account for cognitive tendencies in crafting legal standards and precedents as well as legal education. Finally, it posits that practitioners can maximize their odds of success on motions to exclude expert witnesses through similar awareness of what influences judicial conduct, especially but not limited to cognitive limitations in rendering judgments under time constraints and conditions of uncertainty.

Dedication

To Brendan and Eleanor – you may never want to read this, but I truly hope you understand what it represents. You two are the reason I have chased this dream – and I look forward to seeing you go after yours.

Acknowledgements

A complete list of everyone who had a hand in the completion of this dissertation would likely be longer than the dissertation itself. So I will limit this to acknowledging those who suffered through the last several years rather than my entire life to date.

First and foremost, I have to thank Lilia Teninty – she is my most reliable friend (maybe better friends now that we are not married anymore) and someone who has kept me on track towards completing this seemingly impossible task for so long. My kids, Brendan and Eleanor, kept me inspired in this pursuit and gave me something to focus on/divert my attention to whenever I needed a break.

I never would have made it past the first year at Portland State without an amazing cohort, so I thank each and every one of them, along with the faculty who challenged us and the staff who helped us survive and thrive. Two members of my cohort stood out – Erin Elliot for giving me the push to start presenting at conferences and being amazingly supportive and positive, and Rebecca Craven for putting up with sharing a cubicle with me and always being there to share teaching triumphs or commiserate over setbacks.

My colleagues at Whittier College deserve special mention as well. The belief they showed in me provided the impetus to bring this project to completion. While I am beholden to the faculty, staff and administration as a whole and the denizens of Platner Hall in particular, my Political Science colleagues – Sara Angevine, Fred Bergerson, Joyce Kaufman, Mike McBride and Deborah Norden – have been by turns supportive, inspiring and slightly intimidating on a daily basis. I also owe my actual ability to get this project done to our amazing administrative support Olivia Solis – I would have been

lost without her insights and her steady supply of snacks. I hope my second year is as rewarding as my first, even if it doesn't involve staying quite so late quite so often.

I never would have gotten a position at Whittier if not for getting an adjunct teaching post at Linfield College. For that, I have to thank Bruce Gilley for referring me and Nick Buccola for being an amazing department chair and colleague.

Finally, every member of my committee played a key role in shaping this project. Robert Fountain demystified statistical analysis for me, no small feat given that my last math class was in 1990. Birol Yesilada exposed me to the broader field of decision-making just as I was starting to investigate, and be thoroughly confused by, heuristics and biases. Every time I spoke with Phil Cooper I learned something new about fields ranging from law to sustainable development to national security policy. One day I would like to intimidate, inspire and enlighten my students the way he did me. Christopher Shortell has had the misfortune of being both my boss and my dissertation chair for several years. He has never been less than kind and generous, while demanding that my work is never less than my best. I hope I have lived up to his expectations.

All errors are my own, and hopefully not too common.

Table of Contents

Abstract	i
Dedication	iii
Acknowledgements	iv
List of Tables	viii
List of Figures	ix
Ch. 1 –How Everyday Rulings Can Shine a Light on Judicial Decision Making	1
A. The Ongoing Debate Over Judicial Decision making – What Judges Say They Do, and What Political Scientists Assert is at Work	4
B. The Influence of District Courts	11
C. Expert Witness Admissibility – an Acute Example of Trial Court Influence	14
D. Why This Matters – the Challenge of Judicial Legitimacy in a Polarized Society	21
E. Taking a Close Look at Legal Rulings, Judicial Traits and Institutional Factors – a Multiple Methods Approach	29
F. Ultimately, Judicial Decisions Seem Guided by a Common Set of Norms Rather than Doctrinal Law or Politics – and that has Significant Implications for Theory and Practice	34
Ch. 2 –The Main Models of Judicial Decision Making, how Everyday Rulings may Reflect Them, and the Research Questions this Dissertation Explores	43
A. The Default Legal Model – District Courts Follow Rather Than Make the Law ..	45
B. Behavioral and Attitudinal Models – Judges’ Decisions Reflect Who They Are and Their Policy Interests	60
C. Strategic/Institutional Models of Judicial Decision Making – Judges’ Decisions are Constrained by Circuit Practice and Docket Pressure	76
D. A Cognitive Model of Judicial Decision Making – Judges’ Decisions are a Function of Heuristic Processes	85
E. The Research Questions about Expert Witness Admissibility this Dissertation Explores and the General Research Plan	100
Ch. 3 - Thematic Analysis Suggests Judges Favor Legal Norms Over Strict Application of <i>Daubert/Kumho</i>	108
A. The Qualitative Investigation Process	111
B. Excluding an Expert is an Unusual Step	116
C. Judges Rely on Socially Constructed Standards to Assess Methods and Particular Experts	126
D. Judges’ Rulings Demonstrate Significant Cognitive Simplification	144

E. Decisions to Exclude or Limit Testimony are Largely About Addressing Extreme Cases and Policing Role Boundaries.....	156
Ch. 4 - Quantitative Evidence Suggests Judges' Primary Influences are Legal and Cognitive, but Political Identity and Location Matters at the Margins.....	177
A. The Hypotheses Under Quantitative Investigation and Variables Assessed	180
B. The Quantitative Evidence is Broadly Consistent with the Qualitative Investigation Regarding the Influence of Law and Cognitive Shortcuts	205
C. There is Suggestive Evidence for Multiple Influences on Judicial Decision Making	218
D. Regression Analysis Supports the influence of Law, Cognition and Political Affiliation on Judges	236
Ch. 5 - Conclusion – Seeing Judicial Decision Making as a Mosaic and Plotting a Path Forward	253
A. There are Several Extensions and Refinements of this Study that Deserve Attention.....	256
B. This Study Shows Judicial Decision Making is a Mosaic – and Scholars Should Embrace its Complexity and Nuance	261
C. Policy Makers Should be Mindful of Cognitive Influences in Crafting Legal Doctrine and the Need for Shifting Cultural Norms to Achieve Lasting Changes	269
D. Practitioners can Maximize the Probabilities of Success on Challenges to Expert Witnesses by Emphasizing Norm Violations Rather than Technical Arguments.....	276
E. Judges are, in the end, Human and Complex – even when Deciding Whether to Admit Expert Witnesses.....	280
References.....	283

List of Tables

Table 1: Models of Judicial Decision Making.....	45
Table 2: Themes Identified in Qualitative Analysis	116
Table 3: Case Nature Variables.	191
Table 4: Judges Represented in the Sample.....	192
Table 5: Judicial Appointments by Circuit and Administration	197
Table 6: Court and Judicial Trait Variables.....	200
Table 7: Challenge Type and Success Variables.	203
Table 8: Challenge Outcome/Judicial Conduct Variables.....	204
Table 9: Frequency of Admissions and Exclusions.....	206
Table 10: Frequency and Success Rate by Challenge Type	208
Table 11: Frequency of Outcome by Successful Challenge Type	209
Table 12: Exclusion of Challenged Experts by Circuit	212
Table 13: Restrictions of Expert Testimony by Circuit in Cases Where Expert Not Excluded	214
Table 14: Overall Rates of Judicial Action by Circuit.....	215
Table 15: Rates of Upholding Challenges to Qualifications by Circuit	217
Table 16: Rates of Upholding Challenges to Relevance by Circuit	217
Table 17: Relationship of Partisanship to Exclusion of Experts.....	219
Table 18: Relationship of Partisanship to Restriction of Experts	219
Table 19: Relationship of Partisanship to Overall Judicial Action.....	220
Table 20: Relationship of Partisanship to Exclusion of Experts Controlling for Location, Limited to cases in Seventh and Ninth Circuits.....	222
Table 21: Relationship of Partisanship to Restriction of Experts Controlling for Location, Limited to cases in Seventh and Ninth Circuits.....	223
Table 22: Relationship of Partisanship to Judicial Action Controlling for Location, Limited to cases in Seventh and Ninth Circuits.....	223
Table 23: Impact of Partisanship on Restriction and Exclusion of Experts Controlling for Nature of Party Bringing Challenge	226
Table 24: Impact of Appointment Prior to Completion of <i>Daubert/Kumho</i> , Controlling for Partisanship	231
Table 25: Relationship of Summary Judgment Motion and Restriction of Expert Testimony	232
Table 26: Relationship of Summary Judgment Motion and Exclusion of Expert Testimony	233
Table 27: Effects on Likelihood of Restriction; Logistic Regression Coefficients and Model Summary.....	239
Table 28: Effects on Likelihood of Exclusion; Logistic Regression Coefficients and Model Summary.....	243
Table 29: Effects on Likelihood of Judicial Action; Logistic Regression Coefficients and Model Summary.....	247
Table 30: Probability Differences for Variables Showing Statistically Significant Effects	248
Table 31: Summary of Support of Hypotheses.....	250

List of Figures

Figure 1: A Simplified View of the Legal Model.....	48
Figure 2: A Simplified Representation of the Attitudinal Model.	69
Figure 3: A Simplified Representation of the Strategic/Institutional Model.....	78
Figure 4: Rates of Admission and Exclusion by Circuit.....	212
Figure 5: Restrictions of Expert Testimony by Circuit in Cases Where Expert Not Excluded	214
Figure 6: Overall Rates of Judicial Action by Circuit	216

Ch. 1 –How Everyday Rulings Can Shine a Light on Judicial Decision Making

In the 2016 Presidential election cycle an odd thing happened relating to Federal judges. As expected, there was debate over judicial nominations and Supreme Court nominees in particular, with Hillary Clinton trying to capitalize on the obstructionism of Mitch McConnell in blocking Merrick Garland and Donald Trump promising to appoint a true conservative Justice as opposed to the apparently insufficiently loyal Chief Justice Roberts (Sherman 2016). However, Trump also said that Judge Gonzalo Curiel's denial of a summary judgment motion in a Trump University fraud case was tied to his Hispanic heritage and Trump's hardline stance on undocumented immigrants rather than the law (Totenburg 2017). Many public figures condemned the claim and the case was ultimately settled with Judge Curiel signing off on \$25 million in compensation for thousands of class members (Eder and Medina 2017).

For most commentators it was simply one more momentary outrage in a campaign full of them. For Trump's most ardent supporters this claim of racial solidarity was yet another example of him saying out loud what the "politically correct" were afraid to mention. If nothing else, it surfaced in public a question about how judges make decisions that has occupied scholars in a number of fields— are judges at all levels actually applying the law, or is some other factor or series of factors bearing on their efforts? That is to say, are we a nation of laws at all?

Most of the scholarship in this area has focused on the United States Supreme Court and other appellate bodies, in keeping with the idea that these bodies "make law" to a far greater degree than trial courts (Rachlinski and Wistrich 2017: 205-206). Most work has also tended to focus on legal opinions addressing areas of clear values splits

within American society, such as the scope of protections under the 4th Amendment or sexual harassment (Segal and Spaeth 2002: 317-320; Rachlinski and Wistrich 2017: 206-208). Relatively little work has been centered on trial level courts. Even less work has examined the seemingly routine issues such as evidentiary rulings by Federal District court judges that Trump attacked.

Yet it is these daily occurrences where an average citizen is most likely to encounter the legal system as a witness, party, or juror. While these sorts of everyday rulings may not seem significant, they can have a disproportionate impact on shaping the outcome of disputes (Rowland, Trafficanti and Vernon 2010: 183). Directly, they can set the course of trial or resolve matters short of trial. Indirectly, as in the Trump University case they can set the parameters of negotiation for the settlement that is the overwhelmingly likely result of any legal dispute (Hornby [2007]2013: 113-114).

Analyzing influences on trial court judicial decision in pre-trial evidentiary rulings is the focus of the present study. Specifically, it explores the seemingly mundane area of judicial decisions to admit or reject expert witnesses in Federal civil cases. These decisions present an acute case of potential cross-currents between legal rules and norms, politics and personal views. Judges have been applying a fairly well-defined set of legal standards to the question for over a decade, which should lead to the emergence of near consensus as to approach and outcome (Posner 2008: 43). At the same time, the standards are open to interpretation in many respect and the standard of review for the rulings gives a high degree of discretion to individual judges. These are conditions that support personal traits or politics coming to play in a meaningful way (Segal and Spaeth

2002:92-97; Rowland and Carp 1996: 24-57). These decisions can resolve entire cases, which could incentivize strategic approaches to the actions based on the judge's personal schedule or motivation (Rowland, Trafficanti and Vernon 2010). Finally, the factors for analyzing these rulings ask judges to apply some criteria that are part of every case and some that are novel and unfamiliar, which allows for examination of whether judges engage in any cognitive shortcuts to minimize the demands of the novel criteria. (Bainbridge and Gulati 2002; Beebe 2006).

Ultimately, this study suggests that judges are driven by norms that largely align with both legal doctrine and widely shared values within the legal community. At the same time, this consensus approach does not necessarily reflect what might be an optimal approach to informed decision making or strict adherence to all elements of legal doctrine. Political affiliations seem to matter marginally, but in a secondary and much smaller way than legal precepts. There is also evidence that that where a case originates geographically is as significant to outcomes as the political affiliation of the judge, and that judicial action does follow some expectations of behavioral economics. In short, judicial decision making is a result of a range of influences rather than fitting any one explanatory model.

These findings in turn have some significant implications for both scholarship and legal practice. For the former, it suggests that much more nuance is required in attempting to use models originating in Supreme Court jurisprudence to understand trial court decisions. For the latter it advises selectivity in challenging experts on the grounds their methods rather than framing it as a question of relevance or a violation of the proper

scope of expert testimony. It also provides some empirical evidence for Geyh's (2016) contention that effectuating lasting change to legal standards is often a function of generational shifts among both the legal community and the public at large, rather than immediate and full compliance with new legal doctrine. Finally, it suggests, in an echo of Holmes ([1897]1997) that all policymakers should consider how judges will view their obligations to comply with the standards they impose – as a judge will often take the path of lowest cognitive burden that meets their obligations of compliance.

A. The Ongoing Debate Over Judicial Decision making – What Judges Say They Do, and What Political Scientists Assert is at Work

Western norms of justice place some fairly severe restrictions on the "acceptable" conduct of a judge. The central tenet of the “rule of law” is that laws will be applied to similar situations and similarly situated individuals in the same way, which distinguishes law from power (Hart 1994:159-163; Fuller 1977:81-91). Locke ([1690]2002: 57-59) argued that the possibility of a disinterested arbiter of interpersonal disputes was one of the key benefits of society that could lure man out of the state of nature. In doing so, Locke was building on Aristotle's conception of disinterested judgment as the cornerstone of “corrective justice” (Posner 1995:391).

This requirement of disinterest has been internalized by both the public and the legal profession as a need for the appearance of something approximating pure neutrality (Posner 1995:229-230). Judges are not supposed to be just human beings dealing with messy, human problems. Instead they must aspire to be Blackstone's (1765 Section 3: 69) “living oracles,” “sworn to determine, not according to his [or her] own private judgment, but according to the known laws and customs of the land.” In placing Federal

judges outside of the direct political fray in Article III, the framers of the United States Constitution hoped that they were creating a learned body that would be able to resist both popular whim and the designs of other branches of government (Hamilton, Madison, Jay [1788]2003: 463-471). All of that suggests a narrow acceptable scope for judicial conduct – anything that overtly speaks to personal preference is illegitimate. The law is the law.

This is certainly the view that is embraced in most legal education. As Geyh (2016:81-85) illustrates at some length, the "rule of law" is the central lens of law school. Everyone should look first to the existing legal rules, even in cases seemingly of first impression. Novel or indeterminate questions allow for creativity in the construction and application of authority, but the superior answer in the first instance is always one that can demonstrate how the current position is an extension of existing doctrine. Policy considerations are muted, to be used only where authorized by a lack of clear precedent – and even then, rarely presented with the sort of evidentiary clarity one might see in a public policy proposal such as a cost-effectiveness argument, rather than a broad appeal to policy goals like equality or justice (Posner 2008:239-242).

Judges are, unsurprisingly, reluctant to admit anything but pure legal considerations impact their rulings. Senate confirmation hearings are almost always an exercise in judges decrying any influence other than the law, all the way down (Bandes 2011). In part this is because anything else opens them up to the evergreen and potent political attack of being a "judicial activist" rather than limiting themselves to "the law." (Keck 2014: 252-53). Keck's work shows that even in this polarized era judges rarely

take a truly counter-majoritarian position, at least based on the popular support for their decisions. This in turn updates and validates the arguments of Dahl (1957) that the Supreme Court is a critical but limited policy actor. A judge who admits, as Judge Posner recently did in his "exit interview" after announcing his retirement, to looking for a "sensible resolution" first then checking to see if the law prohibits the outcome is apt to trigger at least a small storm of controversy about the judicial role (Liptak 2017). It is far safer to follow the path of asserting nothing beyond the law is ever considered in reaching their determinations (Segal and Spaeth 2002:48-49; Kozinski [1993]2013: 118-119; Posner 2008: 46,).

There is some evidence that, at least in the American context, the perceived legitimacy of individual actions is enhanced by disassociating legal rulings from anything personal about the author. For example, merely framing Justice Scalia's ruling in *District of Columbia v. Heller* as predicated on the Constitution increased the acceptance of the result (Bailey and Maltzman 2011: 161-164). Moreover this phrasing raised the estimation of Justice Scalia by liberals even though the substantive outcome of the ruling was one they would be expected to dislike. The structured, largely impersonal nature of legal opinions may therefore be a key underpinning of the legitimacy of the legal system, in keeping with the concept of procedural justice (Lind and Tyler 1988; Sevier 2014).

At the same time, there has always been skepticism about whether this conception of legal decision making reflects reality. Like all norms, it assumes a degree of inhuman perfection that is simply at odds with lived experience. Off the bench, the judge has a life, a family, friends, connections and expectations that shape who they are and how they

act (Baum 2006: 25-27). To even obtain a position in the Federal judiciary, moreover, likely means an individual has a very particular set of connections and outlooks, including elite levels of education and at least a passing familiarity with local political issues (Posner 2008: 103). Can we really credit that none of that comes to bear when the judge has to make a decision on a case in front of them? That there isn't at least a moment of human emotion animating the enrobed oracle?

In fact, an overly rigid adherence to dispassionate judging would arguably be counterproductive to the legitimacy of courts in the long term. Judge Posner is fond of pointing out that Aristotle's interpretation of justice did not require a lack of empathy – that is absolute neutrality – but a lack of particularized bias (Posner 2008:87-89). For justice to have any resonance, it must be rooted in the values of a specific community in a specific moment in time. As Jared Diamond has noted, non-Western traditions of dispute resolution often focus on limiting negative inter-group impacts of even accidental harms through apologies and collective compensation based on power dynamics, rather than particularized determinations of individual guilt or liability (Diamond 2012: 87-90). A tribal leader who tried to apply the tenets of comparative negligence in an intertribal dispute in New Guinea would be divisive at best, disastrous at worse – as would be a judge who imposed liability in an American court solely on the ability of one party to pay. It is, for that matter, impossible for a modern observer to fully comprehend how Chief Justice Taney could believe his evaluation of the status of African Americans in the *Dred Scott* decision, or how he could convince himself his ruling would resolve rather than exacerbate the core dispute over slavery roiling America at the time. The justness of

a result is ultimately evaluated through a lens of values that is societally and temporally specific, regardless of claims of absolute truth, abiding principle or “natural law.”

(Morgan and Shinn 2014:9-10; cf. Dworkin 1977, 1986).

Hence, the humanity of judges must inevitably impact judgment. The explicit denial of such influence by judges in confirmation hearings as well as the language of their rulings creates an invitation for observers to level charges of disingenuous conduct and hypocrisy. It is one of the enduring discussions in judicial politics. It resonates in Holmes’ (1881:1) proto-realism and admonition that “[t]he life of the law has not been logic: it has been experience.”

The formal language of law couched in logic and argument is undercut by Frank’s (1936:103) admission of the existence of the “judicial hunch”. The “hunch” hardly conjures an image of a disinterested thinker, acting with fully rational behavior and logic. At the same time, it is in keeping with the modern conception of expertise and intuition espoused by Simon (1955). The value of such intuitive thinking has been developed in the fields of psychology and behavioral economics for half a century (Kahneman 2011). Nonetheless, the insights of these fields have yet to penetrate the consciousness of legal actors to a large degree (Epstein 2016: 2071-72).

Instead, for decades political scientists have been at the vanguard of critiques of the rule of law paradigm. The first compelling evidence challenging this view with solid data was the attitudinal model. As discussed in greater detail in Chapter 2, this model – and in particular the study of the Supreme Court by Segal and Spaeth (2002) - argues that the Supreme Court's decisions are primarily driven by the policy preferences of its

members rather than the law. While never fully embraced by institutional or strategic scholars like Clayton (1999) or Epstein and Knight (1998), it was the dominant model for judicial conduct until fairly recently (Geyh 2016: 47-51). Indeed, its impact is such that it has become something approaching "normal science" in the sense of Kuhn (1996:10) that ideological differences will have an influence on judicial decision making. It is all but required that ideology be included as a variable in any study of judicial decision making in the subfield of judicial politics – a requirement embraced in this dissertation.

From that start, a number of other factors have been identified that seem to influence judicial behavior. (Epstein 2016: 2045-2050). A recent review points to studies showing that in particular areas of that law apparent influences on judicial behavior can be linked to gender, race, and age within the general rubric of "behavioral models." (Rachlinski and Wistrich 2017: 206-209). Institutional scholars have long argued for a greater appreciation of the constraints that other actors and political institutions place on the courts at all levels (Epstein and Knight 1998; Clayton 1999; Geyh 2016). Although those arguments were initially dismissed for lack of empirical evidence by attitudinal scholars, recent work has identified a set of potential constraints that most members of the Supreme Court at least appear to consider in reaching their rulings (Bailey and Maltzman 2011). Some of these fit in the classic rule of law paradigm, like precedent, whereas others appear to be more a matter of institutional or strategic considerations including circuit of origin and the views of the Solicitor General (Bailey and Maltzman 2011: 74-76, 82-87, 104-106, 138-139). Other studies show that at least in clear cases judges tend to feel beholden to follow the law in situations lay

people might be inclined to follow other instincts, suggesting a that most judges see legal matters through an isomorphic lens that favors certain types of outcomes (Kahan, et al. 2016: 391-393; Geyh 2016: 51-60). At the same time, some applications of ideas from behavioral economics and cognitive neuroscience suggest that judges are prone to various biases and heuristics, including groupthink, anchoring and availability biases (Miles and Sunstein 2008, Rachlinski and Wistrich 2017). Thus, we are in an age where it is clear a number of factors can influence judicial decision making and generalizations should be viewed with some degree of skepticism (Geyh 2016: 51-55; Epstein 2016: 2045-2050).

At the same time, it is not clear how great the influence of any of these factors really is. In most studies, the outcome was arguably somewhat overdetermined in the study design (Epstein and Knight 2013). Segal and Spaeth looked at Fourth Amendment cases precisely because there was a split in public opinion on these issues, and essentially confirmed that judges are somewhat reflective of the general population. Similarly, the work showing an influence of gender frequently involves issues such as sexual harassment with a pre-existing split in public values that follows gender lines. (Boyd 2016: 795-796; Rachlinski and Wistrich 2017: 207). Very few published studies even attempt to assess how other models might have fared in explaining the conditions in the same circumstances, or how the conduct of judges compared to the general population.

While this has led some to assert that political science has simply failed to understand the nature of legal decision making (Tamanaha 2009; Edwards and Livermore 2009) it also points to the possibility of a new approach in the field. Instead of a narrow focus on the exceptional case likely to wind up at the Supreme Court, looking at the

everyday world of judges might provide a fresh perspective. This in turn might help explain where and to what extent the rule of law paradigm truly holds in practice and why it sometimes fails to explain the conduct of courts.

In addition to its focus on areas of disagreement, empirical work on decision making has tended to focus on the Supreme Court and appellate bodies (Rachlinski and Wistrich 2017: 206, 209-210). That is understandable given the perceived influence of the Court, but it does not necessarily hold much insight for trial court behavior in light of different caseloads, expectations and conduct. A prior generation of work noted the influence of trial courts on policy making (Carp and Rowland 1983, Cooper 1988, Feeley and Rubin 1998), but consciously focused on exceptional cases and issues. As significant as this work is, it leaves unanswered the question of the influences on judicial behavior in the routine cases that make up the vast majority of legal disputes in the United States. In that context, a number of articles have looked at issues like the influence of doctrine or new Supreme Court rulings, (Dixon and Gill 2002) but most have not looked specifically at how judges actually justified their conduct in real cases, instead conducting quantitative analysis solely based on outcomes. Thus, there are still significant gaps in our understanding about how judges operate in the circumstances most likely to impact the daily lives of individuals in our society – the run of the mill tort or contract dispute, the average litigation over the scope of insurance coverage, or a mass tort case pursued as a class action on behalf of thousands or tens of thousands of individuals.

B. The Influence of District Courts

As the Trump University case illustrates, federal district courts are no longer primarily trial courts, but rather the facilitators of resolutions short of jury trial (Hornby

[2007]2013:113-114). This development is a natural and rational response to the perceived nature of jury trials. The American legal system's use of lay jurors to resolve most trials is often critiqued as making the American system less predictable, costlier and more adversarial than that of European or other developed states (Kagan 2001: 113-117). Limiting the number of trials through earlier engagement and active case management has been stressed for decades, leading to greater numbers of civil cases settling and the vast majority of criminal cases resulting in plea bargains (Boyd 2013: 194). The "disappearing jury trial" has been commented on less as a loss for democracy and Constitutional values than the seemingly inevitable shedding of a nuisance (Hornby [2007]2013: 107-111)).

Regardless of the actual number of trials, though, the trial court system still has a significant impact on the law. The perceived costs and probable outcomes of an eventual trial help drive a significant number of settlements (Bornstein and Robicheaux 2008: 10-11; Kagan 2001: 112-117). Given that the overwhelming number of cases brought in district courts are never appealed, this makes trials and all associated procedures a major shaper of the law as practiced and creates the possibility of distinct local patterns of law (Guthrie, Rachlinski and Wistrich 2007: 4-5). Moreover, there is evidence that district court proceedings frequently play a significant role in shaping policy choices across jurisdictions, as both parties and observers of initial proceedings apply lessons learned to limit potential exposure (Zaring 2004). Therefore, any action by a judge that in turn influences trial outcomes is significant.

One such action judges take is ruling on the admissibility of evidence. The American legal system is often singled out as uniquely party driven in regards to identifying and presenting evidence (Kagan 2001: 105-106). The parties, rather than any central authority, take the primary burden in identifying and investigating the issues that are disputed in the case, compiling the materials and witnesses that relate to those issues and presenting the results to the judge or jury in appropriate circumstances. This level of party control is often asserted to be a key factor as to why litigation in the United States is more complex, expensive and time consuming than in other developed societies

Nonetheless, this leeway is not absolute. The legal system places a number of limitations on what can be presented and thus what jurors are allowed to consider in arriving at a verdict. This includes requirements that the evidence be relevant to the issues at hand, probative of the questions to be resolved by the jury, not unduly prejudicial or inflammatory and similar considerations (Federal Rules of Evidence (FRE) 2014). One of the critical themes in evidence law is that to be admissible evidence must be reliable, in the sense that it is unlikely to be a false statement and is not subject to issues of faulty recollection or memory (Beecher-Monas 2011: 38).

This last requirement of reliability is worth clarifying, as it points to a key distinction between the epistemologies of law and science. In the law, as noted, the key concern of reliability is the ultimate trustworthiness of the statement – can it be used as the basis for establishing an element of a case, under a set of norms developed over time that rely primarily on human experience and some insights from early research in psychiatry. It is from that perspective that judges are required, as discussed below, to

assess the “reliability” of expert testimony. It is not, as used in science, related to an instrument’s ability to generate consistent results. Moreover, at the trial court level the initial admissibility determination is not the same as a validity assessment in science – the judge is, in theory, only determining if the evidence fits the legal standard of reliability. The ultimate judgment of the truth of the statement is left to the finder of fact, which may be the judge or a jury. It thus can mean two very different things for a study like this to attempt to assess the reliability of measurements of judicial determinations of the reliability of expert testimony. Throughout this study, the term is used in its legal sense, but the distinction points up in one small way the different world view of judges and social scientists.

As mentioned above, in a jury case the judge is only making a preliminary determination on questions relating to evidence. The jury remains free to accept or reject the admitted evidence based on its collective assessment of the credibility and weight to afford evidence. Nonetheless, the trial judge circumscribes the universe of facts the jury can consider by application of the rules, case law and their own best assessment of a variety of considerations. The resulting influence is so strong that some authors have semi-seriously asserted that “[e]very jury trial is a bench trial.” (Rowland, Traficanti and Vernon 2010: 183). Yet very little of the empirical work in political science engages with district courts generally or these evidentiary matters in particular.

C. Expert Witness Admissibility – an Acute Example of Trial Court Influence

There are a number of pre-trial rulings that judges make, each with their own potential insights into judicial decision making. Expert witness admissibility, though,

presents a unique set of possibilities. It is an area with a fairly well developed legal doctrine, but with sufficient ambiguity to permit a judge leeway to be influenced by extralegal considerations. Judicial review of this decision is highly deferential, so in theory a judge could be incentivized to follow their policy preferences or simply eliminate the case by eliminating the expert. It is a required task, but one that frequently requires a judge to assess unfamiliar topics, creating an incentive to find ways to minimize the mental burden of new information. At the same time, it isn't an area with a persistent likely split in attitudes along political, gender or racial lines, thus creating a meaningful test of multiple models of judicial behavior.

To appreciate the significance of this task, it is worth exploring what an expert is and how judges came to have a responsibility to assess all expert testimony for reliability. Per Federal Rule of Evidence (FRE) 702, an expert is anyone possessing training, knowledge or experience that would “assist the trier of fact to understand the evidence or to determine a fact in issue.” As the rule suggests, there is no hard and fast dividing line between lay and expert witness, and no boundaries on the scope of potential expertise. They can range from a mechanic who can explain how a faulty repair made a car unsafe to a Ph.D. biochemist explaining how the reagents in a blood glucose monitoring strip work. The common thread is that their expertise will provide a different context or meaning for jurors than the raw information would.

Where a witness is speaking to something that is within their field of expertise and will assist the trier of fact, expert witnesses are free of some of the key constraints that limit the admission of other evidence (FRE 702 *et seq.*; cf. FRE 602, 701). They are

not limited to matters within their personal knowledge but can rely on the words and work of others. They can offer opinions, rather than just factual statements, so long as the opinions rest on their expertise. They can even offer opinions as to the ultimate question in a case, although they generally cannot testify in the form of a statement that explicitly tells the jury how to find on an issue like negligence. In many cases, expert witnesses are the linchpin of a party's presentation on everything from the fact of harm to the causation of harm to the amount of damages that resulted, a way of dramatically weaving together disparate strands of evidence into a complete picture (Vidmer 1995: 171-173).

Empirical studies indicate that expert witness testimony can be a critical element in getting a case to trial and succeeding at trial. Initially, the rejection of expert testimony is strongly correlated with judges terminating matters short of trial, either by summary judgment or directed verdict (Beecher-Monas 2011: 39; Dixon and Gill 2002: 294-296). This occurs frequently where precedent requires scientific evidence of causation such as in most toxic tort cases, but in theory it could happen any time expert testimony is the basis for an element of a claim. Dixon and Gill may well understate the effect of exclusion on case termination, as they only looked at cases in which an opinion included both a ruling on a request to exclude and a motion for summary judgment. It may be that exclusion drives disfavored parties to settle or otherwise leads to a resolution even where it is not immediate. Unfortunately, barring reviewing the docket of all cases in which an expert is excluded it is impossible to assess the full impact exclusion has on case resolution.

Moreover, the fact that the expert is even admitted appears to make their testimony more credible in the eyes of jurors than if the testimony is presented without such an endorsement (Schweizer and Saks 2009: 11-13). This suggests that the jury is taking on a clear cue from the judge that the material must be significant and at some level true. While it might seem counterintuitive, such testimony appears to be most impactful when offered on behalf of one party or the other in a way that suggests how the testimony should be applied to resolve the case, rather than neutral information necessary to understand the case (Brekke, et al. 1991: 468-470). That is, Americans appear to be so attuned to the adversarial framework of a legal dispute that when material is presented without a clear sense of who it benefits, it is treated less as evidence and more as background information.

Because of these factors, there has long been concern that jurors may be swayed by convincing but misleading expert testimony (Schauer and Spellman 2013: 13-14). The basic idea is that jurors are apt to overvalue testimony presented by someone with a degree of expertise. There is an experimental basis for this fear – people generally assess experts, like other witnesses, based on issues like appearance and demeanor and background as much as the actual content of their testimony (Vidmer 1995: 122-123, 150, 171-173). Given that most people are not trained in the scientific method or even basic critical analysis, they unsurprisingly often fail to identify or understand the value of validity problems with scientific evidence (McAuliff, Kovera and Nunez 2009: 252-253; McAuliff and Duckworth 2010: 495-497; Levitan 2017: xiii-xxii). Moreover, trying to provide scientific education in the guise of expert testimony is difficult - using additional

witnesses to point out problems in testimony actually tends to result in test subjects distrusting both sides' experts rather than causing them to accept valid and reject invalid testimony (Levett and Kovera 2007: 369-370). Thus, there does seem to be a need for some measure of screening to minimize the potential for expert testimony to mislead.

By 1993, a conflict had arisen over how judges are supposed to assess expert witnesses. Many Federal courts at least nominally used a simple but controversial standard – if the expert's testimony conformed to commonly accepted practices in their field it should generally be admitted (Beecher-Monas 2011: 34). Review of district court decisions suggest the standard was not applied all that frequently, but that it was decisive of the outcome when it was applied (Dixon and Gill 2002: 286). This test arguably conflicted with the text of the Federal Rules of Evidence, as it required no effort on the part of the judge to independently assess the proposed testimony. In *Daubert v. Merrill Dow Pharmaceuticals* (1993), the Supreme Court rejected this standard as too cramped in the context of scientific evidence (506 U.S 579, 589). Instead of simply looking to issues of acceptance, courts were supposed to employ a multi-factored approach to assess if the testimony met with standards of scientific rigor in which acceptance was only one factor. Although it asserted that it was creating only a non-exclusive list, most of the considerations specifically mentioned in the opinion were derived from a Popperian conception of science, including falsifiability, error rate/replicability, the use of standards and the subjection of results to peer review (*Daubert* , 506 U.S. at 593-594; Jasanoff 1995: 63).

Subsequently, in *General Electric Co. v. Joiner* (1997) the Court found that a trial court's consideration of the admissibility of expert testimony should be reviewed under an abuse of discretion standard (522 U.S. 136, 142-143). This essentially conceded that such judgments were too particularistic for reviewing courts to assess accurately except in extreme cases. Adding to the complexity of the situation, in *Kumho Tire Co. v. Carmichael* (1999) the Court clarified that all proposed expert testimony should be analyzed for intellectual rigor, although the exact factors that might be applied would vary depending on the nature of the testimony (526 U.S. 137 at 146, 150). Together, these cases are referred to as the "*Daubert trilogy*", the substantive standards are sometimes referred to as *Daubert/Kumho* standards, and collectively reflect the latest substantive word from the Supreme Court as to how trial courts are to assess the admissibility of proffered expert witness testimony.

The *Daubert* trilogy, and *Daubert* in particular, has sparked several debates in a variety of fields about the interaction of science and the law. Some have hailed it as helping to beat back a crest of "junk science", at least initially (Schwartz and Silverman 2006: 226-231). Others see *Daubert* as fundamentally misunderstanding the nature of knowledge generation in science, and generating artificial boundaries based on what counts as "real" science (Jasanoff 1995:130-131; Jasanoff 2008: 128-129). Beyea and Berger (2001: 348-360) assert that *Daubert* actually led to the creation of a sub-species of pseudo-scientific standards employed by courts that are at odds with the expectations of scientists in a number of fields. Efforts at assessing *Daubert's* impact on the presentation of science have varied from accounts that it severely impaired the ability of plaintiffs to

establish causation in toxic tort cases to suggestions that it makes essentially no difference since a similar number of experts testify in trials whether *Daubert* or the prior general acceptance standard is used (Beecher-Monas 2011: 39-48; Helland and Klick 2012: 17-31).

In fact, over twenty years later, it is not clear if the *Daubert/Kumho* standards have resulted in consistent, effective screening of proposed testimony. Prior research has indicated that judges believe these cases have altered their behavior (Gatowski, et al. 2001:443-445). This same work also suggested that judges are generally not very literate about science and technology, and therefore may not be capable of effectively analyzing such testimony. Rather, what they do seem to understand are the concepts of peer-review and general acceptance within a community of experts. This finding is consistent with analysis of court opinions which found general acceptance to be the factor most often noted by judges in assessing evidence (Dixon and Gill 2002: 284-285).

To date, analysis of admissibility rulings has taken one of several forms. There have been a significant number of doctrinal analyses, which use unsystematically selected cases to assert trends in practice. There have also been several assessments of outcome looking at the impact or lack thereof of shifting from *Frye* to *Daubert/Kumho*. There have also been descriptive assessments of the conduct of judges, including some textual analysis of rulings and survey responses. But there has been only a limited effort at assessing the actual decision making process, and that has relied on outcomes rather than the actual text of the rulings in keeping with the bulk of empirical social science on judicial decision making. Thus, there remains multiple levels of mystery about the

admission of experts – we have some indicators of what judges are doing, but far less of an explanation of why.

D. Why This Matters – the Challenge of Judicial Legitimacy in a Polarized Society

This gap in our knowledge and studies suggesting the impact of extra-legal influences on judges takes on special resonance today in light of the polarization of the American electorate and declining trust specifically in the judicial branch (Pew 2014, Jones 2015). In a broad sense polarization is a reversion to a prior, pre-New Deal pattern of behavior, centered on America's evergreen debates over the role of the central government and the scope of individual liberty. However, in another sense it is distinct because of the intervening changes in the demographics of our nation and the political activation of so many formerly latent groups. Just in terms of the sheer diversity of our political actors we are in unprecedented territory. It is possible, therefore, that we may similarly be confronting new territory in terms of the conduct of judicial actors and whether their actions are still generally aligned with previously espoused norms. In a worst-case scenario, we may be leaving behind even the normative claim to be governed by the rule of law, and instead exist in a world where legal outcomes are a matter of the personal and political views of individual actors as they relate to any particular case or group of litigants.

Of course, judges themselves generally reject this assertion. They claim that they follow the rule of law scrupulously, despite their varied approaches to interpretation (Posner 2008: 191-197; Breyer 2010: 82-84). When asked about controversial cases, they will frequently cite how they acted only reluctantly, and only to the extent required by

law (Justice [1992]2013: 46-49). They will assert that they followed the evidence and applied the law, and if that led to a conclusion they found personally problematic they nonetheless upheld it (Edwards and Livermore 2009: 1897-1898; Breyer 2008: 83-84).

Judges though, have generally not presented much empirical evidence to support these claims. They can point to anecdotal examples, or the common values they assert to share with fellow judges but defining how this can be demonstrated more broadly has been difficult. Political scientists, on the other hand, have had little difficulty finding circumstances in which political leanings, gender and other factors seem to correlate with the conduct of members of the judiciary (Rachlinski and Wistrich 2017: 205-209). This has been especially acute at the Supreme Court, and the focus of 40 years of “judicial attitude” scholarship. The case at the Supreme Court has at times seemed compelling enough to lead even appellate court judges to label it a uniquely “political” court while asserting that the judiciary generally is not subject to this criticism (Posner 2008).

At one level, it would be naïve to assert that courts are divorced from politics entirely. Far too much research has shown that the Supreme Court, for example, often renders opinions that reflect the political values of its members – particularly on politically salient issues like the scope of right to handgun ownership under the Second Amendment or the meaning of “unreasonable” search and seizure under the Fourth (Segal and Spaeth 2002: 324-326). Not to mention such seemingly outcome-oriented ruling as the effective awarding of the presidency in *Gore v. Bush*. There is also a substantial body of evidence that the political makeup of appellate panels can have an appreciable impact on rulings (Miles and Sunstein 2008: 838-840, 846-850).

Moreover, the gap between the ostensible rational and legal conclusions of judges and their seeming politics has been a feature of the United States Supreme Court nearly from its inception. *Marbury v. Madison* is widely credited for establishing the doctrine of judicial review and beginning the process of giving the Court a meaningful part to play in governance. Generations of law students around the world have read Chief Justice Marshall's cogent analysis of the nature of the writ of mandamus and how a commission is effective upon completion, with delivery a merely ministerial act, followed by his almost apologetic explanation that the Court is powerless to act due to the language of Article III. The case has been held up as a masterstroke of legal reasoning and long-range strategy, with Marshall turning a short-term defeat for his Federalist faction (in the loss of judicial postings) into a lasting victory for the Federalist vision of a robust central government, including the judicial branch.

At the same time, scholars who look beyond the four corners of the opinion have pointed out the backstory of the case presents a far less heroic situation (Epstein and Walker 2013: 68-69; Breyer 2010: 12-21). They note that by modern standards Marshall likely should have recused himself as he was the Secretary of State who had sealed and yet failed to deliver the commissions at issue. They point to the inconvenient fact that Marshall's language overlooked the precedent of the Court in granting prior writs of mandamus despite the supposedly clear limits of Article III. They clarify that despite the articulation of the ruling as if it was the product of a well debated adversarial case the Jefferson administration as well as the Jeffersonian Congress had tried to prevent the Court from taking any action. The administration had in fact shown its contempt for the

Court by not even entering a response to Marbury's application and were unlikely to recognize any adverse ruling. Indeed, this scholarship indicates Congress had openly targeted the Chief Justice for impeachment proceedings if he took an overtly partisan in this case. Thus, while the path Marshall threaded through the thicket of issues was indeed fine, the result was as much a product of contemporary politics as it was about lasting legal issues.

Thus, we should at least expect some political influence in many of the rulings of the Supreme Court. Often though, this can easily be reconciled with their fundamental roles, as they navigate what Hart (1994:145) called the "open texture" of the law and language. However, what has seemed to grow in the last several decades is the perception both in Congress and among the public that courts at all levels are political. A related issue that sometimes surfaces in popular accounts is the perception that diversity on the bench results in inherent biases influencing outcomes. If there is any truth to these beliefs, it would pose a significant question to the legitimacy of the entire court system. Indeed, it would seem to question the value of a judicial system in its entirety.

Interestingly, the challenge to the legitimacy of the legal system line is itself polarized. In the last decade, conservatives have expressed a disproportionate degree of dissatisfaction with the system (Jones 2015). This is so despite the fact the Court's majority has been appointed by Republican Presidents since 1972, and that most scholars consider the Court to have taken a shift in a conservative direction for the last several decades. (Epstein and Walker 2013:41-50). Apparently, the problem has been that that actors on the Court have failed to live up to the ideals of the Republican base, particularly

in the realm of personal liberties (such as gay marriage) and the failure to strike down the entirety of the Affordable Care Act. Again, President Trump provides a good example, having lambasted Chief Justice Roberts as a “nightmare for conservatives.” (Rossoll 2016).

Despite the criticism of President Trump’s specific comments of Judge Curiel and Chief Justice Roberts, his views are in line with a perception that vacancies on courts at all levels are primarily political spoils. While some trace the rise of a politicized view of appointments to the failed nomination of Robert Bork, the last two decades have seen a significant rise in contested nominations at all levels (Kritzer 2007). For most of the 20th century, it was highly unusual for nominations below the Supreme Court to be contentious – most nominees who received a hearing were confirmed nearly unanimously. Instead, the greater fight was either intraparty over nominees such as the resistance of Judiciary Chairman Eastland to appointments of liberal Democrats, especially in the South. (Carp and Rowland 1983). Now obstructionist tactics have been on the rise by both parties and bulwarks favoring consensus, such as the filibuster and blue slip, have come under attack.

The Obama administration's record with nominations illustrates these issues. Initially, even when Democrats had a near filibuster proof majority in the Senate nominees were delayed, with cloture having to be invoked to confirm many nominees (Beth 2017). This practice became routine enough that ultimately Democrats resorted to the “nuclear option” of changing the cloture rules to only require a simple majority for everything except the Supreme Court. Once Republicans obtained a majority after the

2014 elections, the pace of confirmations slowed even further. While the complete refusal to consider the appointment of Merrick Garland for months was the most well-known, it was far from the only example of this behavior. Cumulatively, the delaying efforts led to the expiration of 59 nominations pending on Election Day and ultimately created over 100 open seats in the Federal judiciary when Donald Trump took office (Beth 2017; Collins 2016).

Although the behavior of Senator McConnell and the language of President Trump are outliers, they are not alone in casting a skeptical eye on the courts from all perspectives. When Judge Vaughn Walker came out as gay in 2011, after having ruled in *Perry v. Schwarzenegger* that California's ban on gay marriage (Proposition 8) violated the Fourteenth Amendment, the National Organization for Marriage sought to have the ruling overturned on the grounds of his orientation (Egelko 2014). Specifically, the group alleged that he was hopelessly conflicted as he might one day want to marry, and thus self-interested. Similarly, a broad range of conservative actors sought to have Justices Ginsburg and Kagan recuse themselves from considering *Obergefell v. Hodges* as they had officiated at same sex wedding ceremonies (Wheeler 2015). Oddly, there was no suggestion that the harm gay marriage supposedly posed to heterosexuals disqualified any of the conservative members of the Court based on their marital status.

Liberals are no strangers to linking politics and the courts, of course. *Bush v. Gore*, with its 5-4 partisan split effectively terminating the 2000 election recount is Exhibit A of the link for both the most vocal advocates of the "judicial attitude" model and most rank and file Democrats (Segal and Spaeth 2002: 1-3). Similarly, the choice of

a single judge in Texas to block President Obama's Deferred Action for Parents of Childhood Arrivals (DAPA) on a nationwide basis, which was affirmed by the 5th Circuit and upheld by an evenly split Supreme Court, is invoked as an example of judicial partisanship run amok (Marcus 2015). Or at least it was until judges in Washington, Maryland and Hawaii all used similar logic to impose nation-wide injunctions on the Trump administration's restrictions on immigration from first seven and later six countries in the "travel ban" cases.

The media coverage of these cases further feeds the perception that politics is the driver of the outcome. Almost inevitably, a write-up of a controversial decision will take note of which President appointed the key actors (Marcus 2015). It is frequently used as an implicit suggestion of political motivation, such as when it turned out Judge Hanen – the author of the DAPA injunction – was a George W. Bush appointee. Less commonly, this denotation is used as an implicit designation of the judge as rising above politics – such as when it turned out Judge Robart in Washington, who blocked President Trump's immigration policy, was also appointed by President George W. Bush (Robbins 2015). Notably lacking is any explanation of such features as blue slip holds, which would let readers know these judges were supported by home state Senators, nor is there much in the way of biography in most cases (although it was noted in at least one analysis of Robart's actions that he had worked with a refugee group in the past and had stated "black lives matter" from the bench) (Robbins 2015). While these comments likely result from the desire for media to try and at least hint at a satisfactory narrative to explain the

outcome, it does little to illuminate the actual process the judge or panel went through in arriving at the result.

If this continuous focus on the political nature of the judiciary was wholly accurate, it would raise profound questions for the ongoing viability of the legal system. The resolution of legal disputes in an adversarial process like the one used in the United States necessarily creates results in which one side triumphs and another loses. The overt focus on the combination of bottom line results and political orientation is so reductionist that it suggests there is no broader reason to the law than will. This is unfortunate not only as it goes against the goals of an independent judiciary but because it ignores the effort that judges – especially district judges, who have to distill both law and facts from the filings and evidence of the parties under heavy deadlines – make to explain themselves. The actual rulings of district courts, when considered at all by the media, are summarized at the most general of levels. The focus is almost inevitably on the outcome and impact of the ruling rather than the logic employed to get there, although occasionally the sheer length of an especially voluminous ruling might be included as point of interest. What the casual reader is left with is the suggestion that politics is at least as important as any legal issue.

The collective effect of this coverage is to suggest that judges are just politicians in robes at all levels. This is a claim not even the most ardent “judicial attitude” scholarship has made. Even that body of work argued only that a combination of factors unique to the Supreme Court essentially made it a political body on matters of high political salience on issues with a clear liberal to conservative uniaxial spacing (Segal and

Spaeth 2002: 86-97). Thus, scholarship that takes what judges actually say seriously and documents how their conduct aligns with or deviates from the norms of legal decision making is vitally necessary.

E. Taking a Close Look at Legal Rulings, Judicial Traits and Institutional Factors – a Multiple Methods Approach

One common point in media coverage of courts and a significant portion of the empirical political science literature has been an emphasis on results in areas that are under significant debate within society. In both cases, there has been little credence given to the rationales provided by judges, instead of looking at outcomes in terms of an independent variable such as political leaning, gender, or age (Rachlinski and Wistrich 2017: 205-210). For example, a common approach is to assess judicial outcomes based on the DW-Nominate scores for the nominating President and any relevant Senators to calculate “judicial common space” scores for each of the judges in question (Giles, Hettinger and Peppers 2001).

This approach is understandable. The media has little interest in covering the mundane work of courts, rather than the broad political issues of the day. Similarly, politically salient issues such as sexual discrimination, individual liberties and criminal sentences are more apt to yield interesting (i.e. publishable) results. Indeed, in a moment of Kuhnian consolidation, it is presumed within the field of judicial politics that we should find significant differences in behavior linked to political or personal identity. Ultimately, though, this reductionism ignores both the efforts judges make and tend to overstate statistically significant but relatively small differences (Buchman 2007: 683-684; Rowland and Carp 1996: 27-54).

It is not a new idea to suggest the overt rationales offered by judges for their rulings deserve more credence. Rowland and Carp (1996) amended their own theory of policy making among district court judges in favor of postulating a recognition of the multiple constraints imposed on trial level judges. This was in part credited to the close case study work of Cooper (1988) which revealed that simply reducing judicial conduct to a singular choice by the judge was misleading. Epstein and Knight (2013: 16-18) similarly argue that there should be at least greater focus on the nearly 40 percent of all Supreme Court cases that result in unanimous rulings. When these cases are factored in along with all other decisions in a term, justices normally thought of as polar opposites actually agree more than they disagree, even though the merits docket of the Supreme Court is almost exclusively populated with the most contentious questions imaginable. Moreover, although identifiable differences on particular issues can be found linked to politics or race at the appellate level, the overwhelming norm remains unanimity regardless of the make-up of a particular panel (Miles and Sunstein 2008: 836-840; Kuersten and Songer 2014: 83).

The current study seeks to further advance our understanding of judicial decision making while being mindful of these concerns. It was designed around taking multiple approaches to assessing judicial decision making in the context of the admissibility of expert witnesses. It was consciously designed to look not just at notable cases, but the everyday workings of courts in several distinct geographic regions. Specifically, every identifiable written opinion addressing a substantive challenge to the admissibility of expert testimony in civil cases from 2010-2015 in nine district courts – three each from

the Ninth, Seventh and Fourth United States Courts of Appeals – were identified. This yielded several hundred cases, which were uploaded into atlas.ti for analysis.

Thematic analysis was then applied to assess this information qualitatively. Unlike directed content analysis, thematic analysis permits but does not require that a researcher begin with a fully established set of codes (Braun and Clarke 2006:88-89). In this study, in light of the work of Dixon and Gill (2002), Merlino, Springer and Sigillo (2011) and Kim, Schlanger, Boyd and Martin (2009), factors such as the nature of the parties, the nature of the case, mentions of the *Daubert* factors or other legal frameworks, and indications of positive or negative commentary on these issues were all initially designated as codes. At the same time, consistent with thematic analysis, items of interest that were identified during review were also coded, permitting flexibility to identify what the data establishes rather than only focusing on prior work. Both during and upon completion of the coding, the results were analyzed for the identification of overall themes that emerged from the process. Ultimately, this search identified likely independent variables for quantitative analysis as well as informing assessments about the likelihood that judges were acting consistently with any of several previously identified patterns of judicial behavior in other contexts. In particular, the cases indicated that legal norms and doctrine dominated the overt basis for these cases, but also revealed a pattern of cognitive shortcuts that minimized the burden of decision making and in so doing potentially result in sub-optimal application of legal doctrine as intended by the Supreme Court in *Daubert/Kumho*.

Additionally, consistent with the majority of recent work in political science on judicial decision making, these decisions were subjected to quantitative analysis employing SPSS software (Rachlinski and Wistrich 2017). This included compiling frequencies, cross-tabular and regression analysis using the outcome of an admission decision relating to a particular expert as the dependent variable. Unlike prior studies that only looked at admission or exclusion, additional assessment was undertaken to account for situations in which the judge decided to restrict the testimony of the witness while admitting the expert. Elements of the framework for analysis including the dependent and several independent variables (admission/restriction/exclusion, party status, nature and timing of challenge, location of case) were derived from thematic analysis. Additional independent variables, such as judicial common space scores, gender, time of appointment, and status, were derived from the publicly available biographical information about the judge.

By applying both qualitative and quantitative analysis in this manner, a richer understanding of the process and the outcome in these decisions was achieved consistent with the tenets of multiple method approaches (Johnson, Onwuegbuzie and Turner 2007). Although the qualitative analysis initially yielded an unwieldy number of independent codes, thematic analysis permitted identification of key common issues that predominated and merited closer review. Ultimately these suggest that judges default in favor of admissibility; that they rely on socially constructed understandings of expertise both in terms of particular categories of evidence and particular witnesses; that they seek cognitive simplification of the decision making process where possible, including

simplifying the overall analysis, substituting more readily comprehensible considerations such as qualifications, adherence to generally accepted practices and relevance for actually assessing methods in detail; generally avoiding intervention where possible in favor of trusting the parties to address issues through the adversarial process; and that a concern is policing the boundaries of acceptable behavior by lawyers and experts to reign in extreme behavior, rather than applying a cognitively costly analysis to most disputes.

As interesting as these insights were, they were a prelude to assessing how well the models that dominate the decision making literature fit this conduct. The answer proved to be at once consistent with what most practicing lawyers would likely expect and yet not entirely consistent with doctrinal law, the existing models or even prior work on expert witnesses. Instead, it seems to be a process that is governed by norms both about judicial conduct and the adversarial system. There is a statistically significant impact of partisanship on decisions to exclude experts, but it explains very little of what starts as a relatively small degree of variability. Where a case is located seems to have at least the same explanatory power, and which arguments are raised that are also predictive of judicial conduct as a whole but not as clearly linked to the extreme act of excluding an expert. At the same time, raising the type challenge at the heart of the *Daubert* trilogy – claiming that there is a methodological problem – is actually negatively associated with judicial action. This in turn suggests that judges are at once resistant to the impact of polarization or politicization, but also deviating from a strict application of the law in a way with some potentially troubling implications.

F. Ultimately, Judicial Decisions Seem Guided by a Common Set of Norms Rather than Doctrinal Law or Politics – and that has Significant Implications for Theory and Practice

As noted, the dominant paradigm for most judicial decision making scholarship has been to look for evidences of differences in places where society itself is most polarized. The record on actually establishing these differences is mixed, but on the whole it can be said that in the right circumstances judges with different characteristics will split to some degree. Unfortunately, the narrative that has resulted has been one that emphasizes that potential without putting it in much context in terms of magnitude or frequency. There also has not been much effort to place what differences do exist among judges in perspective relative to other subjects of decision making.

Hopefully the present study will help fill in some of that missing perspective. Initially, there is strong support for the idea that judges generally reach the same conclusion on challenges to the admissibility of expert witnesses. Roughly four out of every five experts challenged are permitted to testify to some degree. Just from these basic numbers, it can be argued the dominant determinant of the outcome is tied to the internal criteria of legal analysis rather than external factors. Overall this suggests that judges do largely try to follow the legal model, a suggestion previously experimentally posited by Kahan, et al (2016) and theoretically by Baum (2006), among others. Moreover, this degree of outcome agreement is similar to that found in studies of other types of judgment ranging from art to wine (Galanter 1993: 71). In other words, the central promise of the rule of law – that similar cases will be resolved in a similar fashion by a disinterested arbiter – appears to be substantially fulfilled by the system in practice.

Moreover, the distinctions that do emerge are multifaceted. When the scope of inquiry is restricted to whether an expert is excluded entirely, the partisan alignment of the judge's appointment has a small but statistically significant impact on the likely outcome, with judges appointed by Republicans tending to exclude a higher percentage of witnesses than judges appointed by Democrats. However, the location of the challenge seems to matter at least as much in terms of explaining the exclusion rates. Although judges in the Ninth Circuit show a similar relative pattern between judges appointed by Democrats and Republicans, it is less pronounced than amongst judges from the Seventh Circuit. Additionally, Democrat appointed judges in the former instance actually exclude more experts as a percentage of challenges considered than Republican appointed judges in any other group studied. As in real estate, location appears to be as critical as any other factor, with the Ninth Circuit having established precedent that seems to require greater scrutiny than other locations.

Furthermore, if the dependent variable is broadened out to include any judicial action to restrict the expert, then the factors that seem to have the greatest explanatory power are the nature of the challenge brought, whether the judge chose to publish the decision, the location of the case, and the political leanings of judge. At the same time, exclusions and restrictions seem to be linked to different factors. Partisanship shows no statistically significant connection to the latter, whereas particular categories of challenge have the inverse pattern. For example, a challenge to the propriety of testimony – that it is beyond the proper role of the witness relative to the court or jury – is very likely to trigger a restriction on testimony, but it is actually negatively associated with excluding

the expert. Only location is connected to both types of outcomes, and even that varies depending on which locations are being compared. Thus even this fairly routine dispute defies simple categorization.

The high degree of basic agreement between judges that most experts should be allowed to testify supports the legal model – judges are trying to apply the law as they understand it. However, it does not necessarily mean judges are following the *Daubert/Kumho* standards rigorously or optimally. Instead, when looking at the actual justifications provided by judges for their action, what emerges is less analysis of the methods employed by experts and more of an effort to avoid such complications where possible. Rather, the most common justifications for rejecting or restricting testimony seem to be tied to deeper norms associated with the basic conception of a procedurally evenhanded but fundamentally adversarial process. Only extreme conduct appears to trigger the sanction of exclusion. This is a finding consistent with self-reported judicial attitudes regarding appointing expert witnesses independently (Cecil and Willging 1993: 18-21).

Another key finding is that the oft recited *Daubert* factors for evaluating methods – testability/replicability, error rate, peer review/publication and general acceptance – typically collapse into a generalized evaluation of the witnesses’ credentials (their qualifications as experts), occasional reference to available standards as a generalized assessment of the basis for the testimony (the reliability of the testimony) and an assessment of how the testimony relates to the issues in disputer (the relevance of the testimony). This appears to be a result of the flexibility of *Daubert*’s “gatekeeping” task,

as interpreted by *Kumho*, and the sheer breadth of expertise judges are asked to evaluate. To use the metaphor of *Daubert*, there is a gate, but it is generally open so long as the expert (and often the lawyers proposing the expert) have adhered to the broad contours of the schema associated with expert testimony in the American legal system and appear to be within the bounds of a given field. Similarly, while there is not a statistically significant difference in admission rates based on the general context of the motion, there is a consistent pattern that motions raised at the summary judgment stage will result in slightly more exclusions and fewer restrictions than those brought as pre-trial determinations. This suggests judges will invest the cognitive resources necessary to evaluate an expert more closely when there is a potential incentive in the form of case reduction, but avoid engaging in what they see as potentially wasted effort. All of this is fairly clear evidence, along with the adoption of boilerplate legal analysis and often truncated factual assessment, that judges are engaging in cognitive minimization in keeping with findings in other areas of law (Rachlinski and Wistrich 2017; Beebe 2006; Beecher-Monas 2001).

At the same time, the findings of general default to admission and cognitive simplification should not be interpreted as arguing judges are completely passive or asleep at the switch. They do apply a strict *Daubert* analysis on rare occasions, such as where the proposed testimony is so out of bounds that the factors supply multiple grounds on which to reject it. Even less frequently a full-fledged *Daubert* analysis might be used to admit otherwise controversial testimony. Even without applying the *Daubert* factors, judges will also exclude a purported expert who truly lacks the ability to offer meaningful

insight on an issue, who admits (or other evidence reveals) that their proposed testimony is lacking in a meaningful foundation or was the product of undue manipulation, and where there is a plain disconnect between the expert's testimony and the facts of the case. Thus, there is a meaningful application of *Daubert*, consistent with survey findings originally reported by Gatowski, et al (2001: 443-448). It just is not as focused on assessing methodology *per se* as *Daubert* suggests – which itself is not surprising given the same work by Gatowski, et al found judges are generally unfamiliar with the subject of research methods. Indeed, challenges to methods are negatively correlated with both restriction and exclusions of experts.

Additionally, consistent with their role as arbiters of a fair but adversarial process, judges very frequently restrict the testimony of experts even as they are admitted. The basis for these restrictions are often tied to a party's attempt to stretch an expert into testimony that breaches the appropriate scope of testimony in an adversarial system – such as testifying to a legal conclusion (and usurping the role of the judge) or to the credibility of other witnesses (and usurping the role of the jury) or attempting to bring in irrelevant or otherwise inadmissible testimony (and thus violating the rules of the game). Again, this indicates that no matter the nominal parameters of *Daubert*, judges clearly place it within the overall framework of the American legal schema, rather than as a distinct assessment of methodology. At the same time, they seem to undertake this restriction in lieu of exclusion – explaining how a challenge to the propriety of testimony is positively correlated with restrictions but negative correlated with exclusions.

Ultimately, these findings suggest that there is a dominant set of norms, a common schema that judges adhere to and that outcomes are the product of the application of those norms. This suggests two conclusions with somewhat contradictory implications for the ongoing legitimacy of the American legal system. On the one hand, the significant degree of isomorphism in the decision making process suggests a system that is resistant to change or manipulation based on extraneous influences, which in turn suggests that faith in the fundamental procedural fairness of the legal system is justified. In particular, the Ninth Circuit's apparent ability to alter the behavior of some judges through the design architecture established in precedent argues that internal rather than external factors are the most directly impactful means of altering judicial conduct.

At the same time, this very cohesion suggests that certain norms about the adversarial process are so deeply embedded with judges that they lead to cognitive shortcuts and a resistance to full compliance with shifts in doctrine or rules such as a new evidentiary standard. *Daubert/Kumho* on their face challenge judges to screen evidence for scientific and technical rigor, but also to honor the normal precepts of the adversarial process. This in turn often leads judges to defer to industry standards and prior legal determinations of acceptable forms of testimony that can become divorced from a truly rigorous examination of the testimony. Over time this can create a "safe harbor" or "legal science" effect, which in turn can allow poorly grounded testimony in and exclude other valid testimony. (Jasanoff 1995:130-131; Jasanoff 2008: 128-129; Beyea and Berger 2001: 348-360; Beecher-Monas 2011). This tendency, seen in all courts and only marginally reduced in the Ninth Circuit, raises questions about whether and to what

degree district court judges are truly following *Daubert/Kumho* or instead primarily follow the norms of the adversarial system.

The significance of these bedrock norms also suggests that policy advocates would do well to consider if the outcome they seek is consistent with legal norms. For example, there has long been a critique of the high stakes nature of adversarial legalism (Kagan 2001). Any immediate push to shift towards a more collaborative approach, or even the more active case management advocated by Chief Justice Roberts, is likely to meet resistance from the views of current judicial actors. At the same time, as Dahl (1957: 285-291) and Keck (2016: 252-258)¹ note the Supreme Court and the bench as a whole is an evolving body whose conduct is typically in line with the broader sentiment of society. In keeping with Shapiro's (1965: 141-154) application of Lindblom's theory of incremental policy change to judicial decision making, and the findings of Bailey and Maltzman (2011:94-100), and Songer, Sheehan and Haire (2003:125-126), normative changes by judges tend to be more gradual and generational. Ultimately, changes in both doctrinal law and the very image of the role of judges is more akin to a Kuhnian scientific revolution as noted by Geyh (2016:42-43). Thus, the best course of action to eventually achieve lasting legal policy change would likely be an effort at shifting legal education as well as a complement to short term case or legislative advocacy.

¹ Although Keck (2016: 3-4, 198-201) at times places Dahl in the same camp as Rosenberg as claiming the Supreme Court is an ineffective policy-maker or "sideshow" this is an uncharitable assessment of Dahl's (1957: 279) clear intent – "[t]o consider the Supreme Court strictly as a legal institution is to underestimate its significance in the American political system." This is Keck's (2016: 258) essential holding as well- that the court is not a neutral legal body but a source of "bipartisan judicial activism" that serves to mediate contentious political issues in a way that gives people of all political persuasions a sense that "courts regularly protect rights that they hold dear."

As a purely practical matter, this study also has insights for attorneys considering making an objection to an expert witness. A literal approach based on the reading of *Daubert* is likely to be taken skeptically by a judge who believes in the ability of cross-examination, alternative testimony and similar steps to correct any false impressions the jury may have. By comparison, an argument that more clearly explains why the proposed testimony violates basic procedural fairness might succeed in its aim even in the absence of a clear methodological error. Additionally, it might benefit counsel looking to shift a judge's mindset to introduce the empirical evidence that suggests that the vaunted principles of cross-examination and dueling experts do not actually achieve their stated goal of clarification (Cooper and Neuhaus 2000; Levett and Kovera 2008). Again, though, that might be a discussion that is better reserved for a future generation of lawyers and judges than in the heat of a present case.

In the pages that follow, the foregoing will be explored in greater depth. Chapter Two provides a review of models of judicial decision making, including normative frameworks, theoretical influences on decision making and experimental results on decision making both generally and in the context of expert witness testimony. It suggests ways that expert witness admissibility decisions might provide insight on these questions and concludes with an overview of the research questions explored in this dissertation. Chapter Three provides a detailed review of a qualitative analysis of the text of judicial opinions, including research design, results of the thematic analysis and discussion of the implications of the same. Chapter Four reviews a quantitative analysis about the outcomes of expert witness proceedings in light of case and judicial

characteristics, employing a research design informed by both the literature and the qualitative analysis. Chapter Five synthesizes the results of the two analyses, setting forth the implications of these findings relating to judicial decision making theory, judicial policy advocacy and practical implications for legal practitioners.

Ultimately, this study does not claim – and could not – that the prevalence of legal norms is either a net good or that it results in objectively superior judicial rulings. In one sense it is likely enhancing of the legitimacy of court decisions – the law is generally applied the same to similar cases. But in another, it suggests that there is always a potential for the spirit of the law to be subverted whenever innovation or change runs up against established norms within the legal community. If nothing else it suggests much closer attention should be paid to what these norms are, how they change, and how they relate to values and norms in the broader American society. Such an effort will make studying judicial decision making more complicated and nuanced – but should also lead to a more holistic appreciation for the work of judicial officers at all levels.

Ch. 2 –The Main Models of Judicial Decision Making, how Everyday Rulings may Reflect Them, and the Research Questions this Dissertation Explores

Decision making in general, and judicial decision making in particular, has received a significant amount of study in recent decades (Epstein 2016). This has resulted in a number of models and theories that criticize the idea that judicial decision making is normally a matter of a judge applying controlling legal standards to established fact. One prevalent view is that individual judges make decisions in keeping with behavioral traits or attitudes, rather than objective efforts. Another is that strategic or institutional factors may have significant influences on outcomes. Research on cognitive psychology suggests judges, like most decision makers, are subject to cognitive shortcuts or heuristics, particularly when confronted with unfamiliar information. Ultimately, none appears to be the full answer – rather, each provides a piece of a mosaic that is still only partially understood. Whether and in what way this holds true at the district court level as well as the Supreme Court and appellate court levels has yet to be considered thoroughly. By considering elements from each approach, this project aims to produce a more robust understanding of how district court’s make decisions in the context of admissibility as well as more generally.

As noted in Lee Epstein's (2016: 2045) recent review of the field, dozens of different factors have been considered for study in the last several decades. Several academic fields have focused on empirical studies of judicial decision making, although such work by legal scholars (as opposed to political scientists, economists, sociologists and psychologists) is still in a developing stage. No study could help to replicate all of the work in this expanding field, but it is possible to look at how a number of models

apply to a common data set. In particular, it is possible in this setting to test if the decisions of judges on admitting expert witnesses exhibit patterns consistent with the expectation of the legal model, attitudinal model, behavioral model, strategic model, or demonstrate any consistent use of heuristics or similar cognitive shortcuts by judicial actors. For ease of reference, the models that will be addressed in this project are set out in Table 1.

Name	Key Elements	Normative/Empirical/ Experimental	Key References
Legal Model	Law, Facts (Broadly held values in some approaches)	Primarily Normative	Wechsler (1959); Dworkin (1988); Edwards and Livermore (2009); Carter and Burke (2010), Kahan, et al (2016)
Judicial Attitude Model	Judge Policy Interests, Pre- existing individual values, issue cues	Primarily Empirical	Segal and Spaeth (2002); Miles and Sunstein (2008); Songer, Sheehan and Haire (2003)
Behavioral Model	Personal traits/values, issue cues	Primarily Empirical	Davis (1993); Boyd (2013); Boyd (2016)
Strategic/Institutional Model	Judicial policy interests/goals, other institutional actors, incentives/risks for action	Primarily Empirical	Bailey and Maltzman (2011); Epstein and Knight (1998); Clayton (1999);
Cognitive Model	Pre-existing judicial norms and schema, simplification of multi- factored	Primarily Experimental	Beebe (2006); Guthrie, Rachlinski and Wistrich (2007); Rachlinski, Guthrie and

	standards or decision making		Wistrich (2013)
--	------------------------------------	--	-----------------

Table 1: Models of Judicial Decision Making.

A. The Default Legal Model – District Courts Follow Rather Than Make the Law

The core premise of most legal training in the United States is that law is an endeavor focused on two related activities – finding facts and applying law to those facts (Clayton 1999: 17-19). Appellate law in theory removes one of these activities, as the facts are to be taken as found by the court below – although the selection of which facts to emphasize or how to characterize them is frequently determinative to the subsequent interpretation and application of the law (Carter and Burke 2010: 10-12, 28-29). At the trial level, and in particular in the context of controlling evidence, a district court is routinely tasked with making a number of subsidiary factual determinations and then applying the law (Rowland, Traficanti and Vernon 2010: 185-189). The law, though, is often assumed to be applied as received from higher courts under principles of *stare decisis*, rather than based on independent analysis of free standing legal considerations (Carter and Burke 2010). The job of the trial judge is thus seen as sifting between arguments raised by parties looking for the precedent that fits the facts.

In the classic legal model of decision making, this is all that should matter – the parties’ presentations of the law and facts before a disinterested judge (Heise 2002: 839-840). The judge is presumed, in a conceit going back to well before Locke, to lack any incentive to favor either side but rather is a rational, objective user of legal principles (Locke [1690]2002: 40-42). This is a position that accords with judges’ own subjective view of their role, as they commonly report seeing themselves as referees or umpires, and

routinely denounce anyone who asserts they are acting as advocates (Kozinski [1993]2013: 116-118). At the same time, this is a task necessarily imbued with value judgments – it is not absolute neutrality, but impartiality in the sense of not having a vested interest in the outcome or relationship to one of the parties (Carter and Burke 2010; Posner 2008: 83).

The legal model's support is primarily normative in nature, rather than empirically established. Hart (1994: 145), for example, suggests that judges are only permitted to make law in the narrow confines of the “open texture” left by other law makers, and that even then they are bound by procedural safeguards. While Hart's core purpose was to establish criteria for distinguishing law from other forms of power in support of legal positivism, his analysis was not rooted in empirical data on the conduct of actual judges. Similarly, Fuller's (1977) famous assertion of law's inherent morality is linked to procedural safeguards and regularity that distinguishes it from other policy choices. However, as with Hart, Fuller's work rests primarily on logic and some measure of legal doctrine, but not meaningful empirical observation.

Another view of the legal model comes from the legal process school and in particular the work of Herbert Wechsler. Wechsler (1959: 6-10) famously sought to cabin the role of the Supreme Court to one narrowly defined by neutral considerations of law rather than anything as arbitrary as social justice. This in turn led him to argue the Warren Court's ruling in *Brown v. Board of Education* was erroneous both because it focused too narrowly on discrimination in schools rather than attacking the concept of “separate but equal in its entirety” and resting on a too expansive view of legitimate

justifications for the Court to review and strike down state laws (Wechsler 1959: 31-35). Invoking broad concepts of justice rooted in an understanding of the impacts of segregation on society was, to Wechsler, the action of a policy actor rather than a valid exercise in legal decision making, even though he admitted he could develop no alternative that would fit his neutral principles and address the central question of segregation. This again, is a normative view – one roundly critiqued by Posner (1995: 70-80).

Perhaps the most strenuous support in relatively recent jurisprudence can be found in the work of Dworkin (1977, 1988). His work asserts the law is differentiated from other elements of governance because of its grounding in principle. In this construct, following principles consistently should lead to “right” answers even in ambiguous cases. This would in turn elevate the perceived messiness of the adversarial model and make for consistently correct and universal judgments. Dworkin particularly characterizes the judge as being obliged to follow the law as a matter of integrity. (Dworkin 1988: 218-219). That is consistent with how many judges see themselves. (Kozinski [1993]2003: 116-119). Judicial internalization of the idea that the law is not just policy-making by other means, but a principled quest for a just resolution helps explain the hostility so many judges express to the findings of empirical research about judicial decision making. (Edwards and Livermore 2009; Epstein 2016: 2071-2072). It is clearly a widespread aspirational goal, and one that should be explored empirically to the extent possible (Heise 2002).

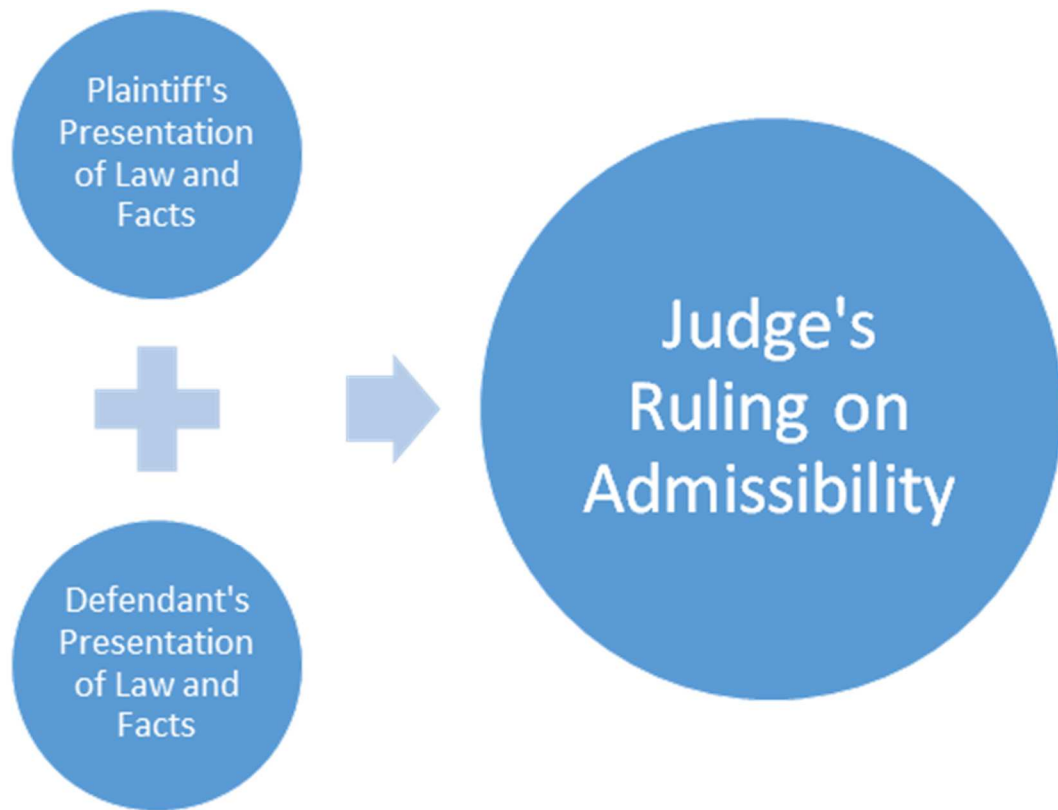


Figure 1: A Simplified View of the Legal Model.

Finding a way to empirically assess the legal model is less obvious. Judges will almost universally say they are applying the law. Appellate judges can frequently point to cases in which they have upheld a decision they see as contrary to their preferred position as evidence of the power of the rule of law. (Edwards and Livermore 2009: 1897-1898; Breyer 2008: 83-84). District court judges will note that they are limited by the potential for reversal on appeal by an appellate bench that often does not share their policy views. (Justice [1992]2013: 46-49; but see Songer, Segal and Cameron (1994: 674-675) (noting possibility of shirking or deviation from precedent where review is deferential or unlikely).

What has been used in other studies of decision making is a consistent set of common outcomes among a particular group or between groups. As Galanter (1993) notes, judge/jury congruence of outcomes was the standard of measurement in the classic Chicago Jury Studies. Similarly consistent outcomes amongst sample populations is generally used in behavioral economics such as the work of Tversky and Kahneman (1979). Thus, perhaps the simplest way to test if the legal model is being used is if judges, as a class, reach similar results in cases applying the same legal standard.

Using consistent outcomes as an indicator of collective following of the legal model, there is significant evidence of the legal model at work in studies of judicial decision making. For example, even though the Supreme Court takes primarily contentious cases where a split of authority is perceived to exist, it has historically reached unanimous decisions in nearly 40 percent of its opinions (Epstein and Knight 2013: 16-17). As a function of overall voting, Justices who are seen as having broadly divergent views nonetheless appear to have some sizeable areas of agreement. Thus, the liberal (at least for the Supreme Court) Ruth Bader Ginsburg can agree in whole or part with colleagues Clarence Thomas and Samuel Alito no less than 62% of the time (Bhatia 2016: 28). At a minimum the legal model can be asserted to be the default model for judicial decision making and the one with perhaps the greatest raw explanatory power.

This trait of uniformity is even more pronounced at the intermediate appellate level in the Federal court system. There, courts are generally facing appeals "of right", meaning the court has to take them and render a verdict. This means the courts have a vastly expanded docket relative to the Supreme Court, handling thousands of cases.

Unlike the Supreme Court, consensus is by far the norm in this level of proceeding with 90 percent of all cases being resolved unanimously (Kuersten and Songer 2014: 83).

This fact is one of several things that leads jurists to argue the legal model is a valid explanation for legal decision making (Tamanaha 2009; Edwards and Livermore 2009).

It should be noted in this context that uniformity is generally measured at the intra-panel, and hence intra-circuit level. It does not necessarily mean one would find a 90 percent congruence on all similar cases across circuits owing to variations in circuit level precedent. As Songer, Sheehan and Haire (2003: 125-126) demonstrate, one of the key sources of change in American legal doctrine within the Federal system are circuit level distinctions that have yet to be resolved by the Supreme Court. Yet if judges subject to these differentiated expectations are reaching decisions that are in line with controlling circuit authority, this would be consistent with the legal model. As discussed below, circuit level developments may also trigger conduct that is consistent with strategic models. For example, if a circuit, without changing the substantive law, increases the sanction for judicial error, a judge may be more cautious in applying the law out of strategic concerns while still applying it in a manner consistent with the legal model.

Regardless of the impact of circuit law, it is less evident if district courts have a similar degree of internal consistency as appellate panels. On the one hand, the lack of salience of most cases to any individual judge and the sheer time cost of unique action might lead to a trend towards conformity (Boyd 2016: 790-791, 795-796.; Taha 2004). At the same time, the relative infrequency of appellate review, the deference given to

lower court decisions in many aspects of appellate review and the isolation of judges all at least create the possibility for significant judicial variation at the trial level (Rowland, Trafficanti and Vernon 2010: 180-182). However, it seems unlikely that appellate decisions would be able to exhibit such a high degree of unanimity, or such a low appeal rate, if lower courts were not in the main consistent as well.

One possible explanation for this degree of consistent outcome in legal settings is the power of social pressure and role conformity. One of the key needs of almost all humans is for the respect of their colleagues, friends and family, what is sometimes referred to as the audience for behavior (Baum 2006: 25-27). For judges, two critical audiences are their judicial colleagues and the lawyers who practice in front of them (Baum 2006: 50-53, 99-100). Law school is largely isomorphic, with an emphasis on case method instruction and forms of reasoning in keeping with the legal model (Epstein 1981: 4-5; Clayton 1999:17-19). Judges are further led by experiences as practicing attorneys to think of their roles in very particular ways (Kozinski [1993]2013). Moreover, either in law school or in practice, future judges also become aware of the fact that the law is not self-enforcing, but that the legal system relies on legitimacy in the eyes of the general public as well as the support of other branches of government to carry out judicial statements (Breyer 2010). The process of judicial nomination and confirmation all but ensures the most radical elements of the bar are further eliminated from the pool of potential decision makers (Posner 2008). Thus, lawyers generally and judges in particular are likely going to be strongly influenced in their official behavior by what they see as the limits of legal conduct.

Judge Kozinski's ([1993]2013:116-120) comments in this regard are informative. During his tenure Judge Kozinski was seen as one of the most conservative appellate judges in the country, which is all the more remarkable given that he sat on the Ninth Circuit Court of Appeals, generally perceived to be the most liberal in its leanings. Judge Kozinski insists that judicial attitude and behavioral models are wrong, and that he adheres to applying the laws to the fact. But even he admits that his work is heavily influenced by norms of acceptable behavior. In particular, he worries about what his clerks and colleagues will think of his efforts, and that these serve as constraints on his conduct – he must explain his conduct in terms they will accept (i.e. the language of law) rather than preferences. He also argues that the necessity to confront precedent, coupled with self-reflection, will generally constrain a judge from acting in a manner that solely reflects their personal interests or behavioral inclinations. While a judge might have a preferred outcome, he argues that these constraints coupled with critical self-reflection will serve to keep the judge's behavior within appropriate bounds.

These constraints have a particular effect within the American legal community because the membership has a shared set of norms about acceptable tools and methods of legal interpretation. The fact that only a handful of tools are generally seen as legitimate basis for statutory or Constitutional interpretation, for example, tends to define the scope of discussion judges employ and in turn bound the potential outcomes (Posner 2008: 191-197; Breyer 2010: 82-84). While judges can and do differ on how to apply precedent or when to give weight to legislative intent, they do not propose a roulette wheel or a duel as an appropriate resolution mechanism. If anything, some jurists find the socialization of

lawyers too constricting, as parties fail to address the implications of their arguments and rather stress purely legal arguments (Posner 2008: 239-242; Breyer 2010:76-87). At the same time, this agreement as to an abstract process does not mean absolute uniformity of outcome, as deciding how to apply the tools of interpretation or weigh potential consequences is still a matter of opinion.

The legal model generally has not been empirically studied using actual case results. Segal and Spaeth (2002:48-85) are typical in setting it forth as a normative standard, then proceeding to show that their attitudinal model more appropriately explains Supreme Court divergence. Legal scholars and jurists like Justice Kagan defend the law as the entirety of their focus, but generally do not seek to establish the scope and impact of the idea through data driven analysis (Epstein 2016: 2071-2072).

Kahan, et al (2016: 380-384, 391-397, 410-414) asserts to demonstrate the importance of the legal model experimentally, but there are some significant questions about this research design. In particular, it was intentionally structured to signal to judges the correct way to answer the issue under the law while providing a compelling factual narrative to tempt non-judges into coming to a different conclusion and a grammatical ambiguity that might confuse lay observers. While the experiment revealed judges (or at least three-quarters of them) are more mindful of the need to follow clear statements of law than the general public, it does not follow that judges would still do so if the legal standards provided them with greater leeway the way they often do in real cases.

Despite this gap, many examinations of the impact of *Daubert* and the doctrine of experts appear to assume that the legal model holds at some level. Merlino, Springer and

Sigillo (2011:2-5)) is a case in point. This study attempts to identify the social construction of particular forms of evidence in the law. It does so by content analysis of 100 cases in each of 4 disciplines of proposed expert testimony – psychological/psychiatric, toxicology, damages and engineering, with no greater specificity. The somewhat cryptic selection discussion asserts these are the “most common” types of evidence from 1993 to 2010 in LEXIS, without stating how this assessment was made. The authors then proceed to assess the factors present in these rulings as a collective and how various factors correlated to admission or rejection.

On the one hand, this largely descriptive effort does not even acknowledge the possibility of circuit level, ideological, behavioral or other variation in treatment by judges. It assumes, in essence, that judges are applying the legal model and then examines which elements of the legal construct are favored by judges as a collective. In doing so, moreover, the authors appear to have chosen evidence issues that were highly debated in the immediate aftermath of *Daubert* such as toxicology but which were not necessarily representative of the vast majority of expert testimony. Indeed, in the present study none of the specific disciplines of psychology, toxicology or engineering were very prominent while damage testimony is a general category rather than a discipline of study or testimony. Thus it is neither representative of how judges treat these questions generally nor does it provide insight into why judges arrive at their conclusions.

Dixon and Gill (2002:256-257) similarly focus on how judges handle *Daubert* challenges rather than what might motivate their decision making. The focus of this study was to evaluate if there was a doctrinal shift in the wake of *Daubert*. While the

authors could show that judges seemed to be facing more challenges to evidence and issuing more complex analyses (264-267), there was no effort made to assess if the judges were influenced by anything other than the change in doctrine and the response of parties to that change. Again, the assumption is that the behavior conforms to the legal model, rather than assessing how it influences judicial rulings relative to other cases.

Other studies in this area are arguably even less informative about how judges actually handle these matters and whether anything approximating the legal model is being followed. Helland and Klick (2012) analyze the impact of *Daubert* by looking at overall rates of expert testimony before and after adoption of the standard in many states. That is, they look to the raw numbers of experts in cases, before and after *Daubert*, and conclude the law did not have a significant impact on behavior. This says nothing about how judges are assessing the evidence. Moreover by looking at the number of experts who testify at trial it ignores how *Daubert/Kumho* rulings can modify the viability of a parties' case and lead to settlement or resolution short of trial. While critical of the impact of *Daubert/Kumho* within the context of trial complexity, it does not illustrate how judges evaluate the law or what factors go into those decisions.

Similarly, Cheng and Yoon (2005) look at removal rates from state courts, with whether the state court has adopted *Daubert* as the independent variable. The authors assert that removal rates in turn suggest whether defense lawyers, as a matter of revealed preference, see *Daubert* as a benefit to their positions. Although this suggests something about the perceived actions of judges in response to changes in legal doctrine it says very little about what judges are actually basing their decisions on.

In fact, only one study appears to have tried to use *Daubert/Kumho* for insight into any element of the modes of judicial decision making rather than doctrinal impact. Ironically, though, Buchman (2007) arguably did not give enough credence to the potential impact of the legal model. Buchman (2007:680 and n. 10) asserts to identify the entire population of tort cases involving expert witness issues from 1983 to 2003. His analysis did reveal a statistically significant link between ideology and admissibility, and no impact from the potential of reversal on appeal based on ideological conflict (Buchman 2007: 683-687). However, in the first instance he did not look for circuit level distinctions in legal doctrine or in fact any level of circuit level differentiation amongst judges despite (as shown by both Dixon and Gill (2002) and this study) the potential for such variation. This is a key oversight in light of the work of Songer, Sheehan and Haire (2003: 125-126) on the impact of circuit level distinctions in practice.

Buchman's data selection protocol may also have led him to accidentally oversample non-representative cases. In attempting to identify tort cases, Buchman reports that he used two search streams, each of which included the word "tort." Due to the Boolean structure of LEXIS' default search engine, this will identify cases in which the term "tort" appears in the plain text, but most cases that sound in tort do not employ that term as a descriptor. This in turn makes his sample of only 705 rulings over 20 years suspect. Additionally, the cases that would be included in such a search would embrace "mass torts" and "toxic torts", which were high saliency issues amongst the "tort reform" movement during the time frame of his study. (Kagan 2001: 147-150).

This oversight in turn likely explains the strikingly low rate of admission reported by Buchman (2007:681). Both *Daubert* itself and the Advisory Committee Notes to Federal Rule of Evidence 702 suggest that admission should be the rule rather than the exception, something Buchman does not mention. Yet he reports an overall admission rate of just under 44 percent despite the fact that over 80 percent of the cases applied *Daubert*. That would be a remarkable indication judges are ignoring the precepts of law if it was borne out by the evidence. However, the likelihood that there are a high proportion of mass or toxic tort cases may also mean a large number of toxicologists. As Merlino, Springer and Sigillo's (2011: 9-15) results show this study evidence was rejected 78% of the time, a far higher rate than other forms of evidence. That rate in turn is consistent with findings that after *Daubert*, the Ninth Circuit's doubling of risk standard for such evidence and its widespread adoption created a unique barrier to toxicology testimony. (Beyea and Berger 2001: 348-350; Beecher-Monas 2011: 41-43). Thus, Buchman's findings may speak to an anomalous subset of expert admissibility rulings rather than the broader category.

Beyond the possible oversampling of toxicologists there is reason to suspect that judges show much more deference on the question of admitting experts than reported by Buchman. Under FRE 706, judges are permitted to appoint expert witnesses either at a party's request or on their own. This is a fairly common approach in foreign legal systems applying the "inquisitorial" model of decision making. (Kagan 2001: 107). In the United States, however, many judges see that as intervening in the adversarial process to an unwarranted degree in anything but exceptional cases (Cecil and Willigang 1993:

18-21). Instead, the expectation is that the parties will curb any excesses or overreach through rebuttal and cross-examination (*Eastman Chemical Company v. Plastipure, Inc.*, 775 F.3d 230 (5th Cir. 2015) (noting with approval trial court's refusal to intervene in a "battle of the experts")).

Even when a circuit court specifically implores a district court to consider appointing an expert witness to sort out complex claims the judge may be reluctant to do so. In a remarkable case, the Seventh Circuit noted that establishing a conspiracy in the pricing of corn syrup had devolved into multiple rounds of dueling regression analysis and debate over the finer points of statistical technique including missing variables, multicollinearity and ways of controlling for it. *In re High Fructose Corn Syrup Antitrust Litigation*, 295 F.3d 651 (7th Cir. 2002). The court went so far as to say that on remand

“[w] recommend that the district judge use the power that Rule 706 of the Federal Rules of Evidence expressly confers upon him to appoint his own expert witness, rather than leave himself and the jury completely at the mercy of the parties' warring experts. . . . The judge and jurors may not understand the neutral expert perfectly but at least they will know that he has no axe to grind, and so, to a degree anyway, they will be able to take his testimony on faith.”

Id., 295 F.3d at 665.

Yet on remand, the lower court still chose not to appoint an expert, relying on the ability of the lawyers to help navigate the jury through a thicket of conflicting and highly technical evidence (Federal Evidence Review 2013).

This incident not only high-lights how much trial judges value their role as disinterested observers, it points to a possible means to operationalize the question of whether judges follow the legal model in deciding whether to admit expert witnesses. If

the normative model holds, then it should follow that judges would reach similar results throughout the nation in light of the well-developed set of Supreme Court precedents from *Daubert* to *Kumho* and the revisions to Federal Rule of Evidence 702 that codified the cases. In particular, the exhortations in *Daubert/Kumho* and the Advisory Committee notes to Rule 702 that admission should be the default should lead judges to err on the side of admission in all but the most extreme cases. There would be expected to be some variation based on the objective strength of the case presented and the skill of the parties presenting matters. Additionally, it is still likely that there would be some variation as there is almost always a residual disagreement between parties' judging the same circumstances (Galanter 1993: 70-71; Heise 2002: 840). Nonetheless, a consistent treatment of experts that followed the guidance of *Daubert/Kumho* and Rule 702 would seem to validate the legal model.

In sum, the actual impact of legal doctrine on judicial behavior in this area remains unclear. Several studies assume it matters, whereas others do not seem to want to acknowledge the possibility. This study tries to establish if we can say there is in fact a general tendency amongst judges related to the admission of expert testimony without assuming its existence. In particular, it looks to overall admission rates within and across circuits to identify whether judges are following the exhortation of the Supreme Court and Rule 702 to default in favor of admissibility. If they are doing so with similar patterns across circuit lines, despite differences in gender, age, political attitudes and strategic opportunities and constraints, this will provide significant evidence to validate the legal model. Moreover, by assessing both this general pattern and then analyzing any

variations in terms of behavioral, attitudinal and strategic models of judicial decision making it should be possible to develop a more complete picture of how all of these influences interact and shape judicial conduct.

B. Behavioral and Attitudinal Models – Judges’ Decisions Reflect Who They Are and Their Policy Interests

The components of the legal model are strongly defended by most lawyers and judges as critical components of what Geyh (2016:2) calls the "rule of law paradigm." However, it is far from a complete explanation of judicial behavior. As noted, disagreements are common between judges, even when applying identical frameworks. The potential for splits on legal issues is one of the key criteria the Supreme Court uses in selecting cases from lower courts. Moreover, as shown by the behavior of the Supreme Court itself it is clearly possible for judges steeped in the same education, applying the same rules to the same facts to arrive at diametrically different conclusions. There are clearly other influences worth studying if the goal is to explain the full range of judicial behavior rather than just a default rule.

The legal model has been attacked as a myth by generations of scholars in law, politics and psychology. Oliver Wendell Holmes, Jr.’s ([1897]1997: 994) seminal assessment that “[t]he prophecies of what the courts will do in fact, and nothing more pretentious, are what I mean by the law” is one of the earliest “modern” attacks on this idea of judges as dispassionate arbitrators. This viewpoint was taken up by the realists in the 1920s and 30s, notably in Frank’s *Law and the Modern Mind* (1936). The legal model has been widely rejected by political scientists, at least as a sufficient explanation

of behavior, and other fields of study have similarly looked with skepticism on the concept (Epstein 2016).

Investigators from several fields have begun to study judicial conduct in the last several years, with a multitude of new models arising to attempt to explain judicial conduct (Epstein 2016: 2040-2050). However, there are two similar but distinct approaches that focus on the individual characteristics of judicial actors that have an outsized place in the literature – the behavioral and attitudinal models (Heise 2002: 833-839). Both began with studies of appellate courts, and neither has proven capable of providing a full explanation of judicial conduct. Nonetheless, given their prominence and the evidence that has been generated to date, they merit consideration as potential explanations of judicial conduct at all levels and were included in the present study for that reason.

Although both the behavioral model and attitudinal model assert that non-legal factors can explain judicial conduct, they vary in their focus. The behavioral model asserts that as with many other aspects of behavior judicial conduct is shaped by personal traits and characteristics that individuals cannot alter, such as gender, ethnicity and training (Heise 2002: 833-836). The attitudinal model, by comparison, asserts that what is more important than these individual traits are the attitudes a judge holds towards particular issues and types of parties (Heise 2002: 836-839). Attitudes can be correlated with behavioral traits in some instances, but in the latter model such unalterable traits are generally not considered. Under both views, it should be possible to see consistent variations in judicial behavior associated with these personal characteristics.

1. Behavioral Traits Generally and Their Potential Relevance to Admissibility Determinations

Behavioralism is arguably simpler to comprehend. It operates on the assumption that socialized experiences associated with the possession of certain traits will in turn be reflected in behavior (Heise 2002: 833). Thus, men will behave in a distinct manner from women, former prosecutors differently from former defense attorneys and even different law school backgrounds resulting in different behavior. The key is identifying significant traits - and as Epstein (2016: 2047-49) notes, numerous different potentially salient traits have been identified ranging from gender to race to the prior work experiences of the judge. This is also, in a very crude manner, the sort of tendency that candidate Trump was referencing when he accused Judge Curiel of bias.

To date, the application of behavioralism to judicial behavior has been uneven (Heise 2002: 834). In many instances, no discernible pattern has been identified (Boyd 2016: 790-791). Again, this may owe to the sort of isomorphism and role conformity that the shared experience of legal training and practice produces in judicial officers. At the same time, in particular circumstances there do seem to be noteworthy variations. For example, female judges seem more likely to support claimants of employment discrimination (Davis 1993:131-132). This is true at both the appellate and trial level (Davis 1993: 131-132; Boyd 2016: 793-796). Similarly, trial judges who are racial minorities tend to sustain EEOC enforcement cases against motions for summary judgment more frequently than their white counterparts (Boyd 2016: 793-796). Experience as a prosecutor or defense attorney appear to lead judges to differ from their

colleagues in certain settings as well (Heise 2002: 835). Thus, behavioral explanations cannot be rejected out of hand.

The primary evidence for behavioral models tends to lie in settings where there is a salient link between the characteristic under investigation and the subject matter of the case. That is, it is more likely to see behavioral correlations with gender on so-called "women's issues" like gender discrimination, and correlations with race on issues raising questions of salience for racial minorities like racial discrimination (Boyd 2016: 790-791, 793-796). Boyd's recent work reflects this tendency. Boyd looked specifically at whether trial judges in the federal system would maintain a discrimination suit brought by the EEOC against a motion for summary judgment. She found that female judges were more likely than male judges to support claims involving gender discrimination, and that minority judges were more likely to reject motions for summary judgment in matters alleging racial discrimination than their white counterparts.

While these findings certainly can be seen as supporting behavioral impacts, they also point to the limits of the theory. While Boyd's findings appear to be significant, the need to so limit the data to find such evidence suggests it is a marginal effect overall. Additionally, the data set used by Boyd ended with motions decided in 2006, thus significantly limiting the number of female and minority judges relative to their current percentage of the Federal judiciary in a way that could skew the results. (Epstein 2016: 2048; Boyd 2016: 788 and 796 n.1). Although the study controlled for partisanship, it did not appear to control for the timing of appointments or length of tenure of judges. Given that Presidents Reagan and George H. W. Bush appointed far fewer female and

minority judges than President Clinton or even President George W. Bush (Boyd 2016: 796 n.1), this finding could be an artefact of the shift in societal values over time as much as it is a function of gender or race.

Indeed, one issue with using behavioral traits in the context of judicial decision making is that the judicial population has been undergoing steady demographic shifts. On the one hand, racial minorities and women remain underrepresented on the bench at all levels (Boyd 2016: 788-789). On the other, diversity has been increasing steadily. Women in particular have made tremendous progress in their appointment to the bench in the last three decades, although they still only represent roughly one third of all judicial officers in the Federal system – and that is not distributed evenly geographically. (Epstein 2016: 2048). Barack Obama in particular significantly increased the percentage of female appointments, representing over forty percent of his more than three hundred appointments. He also increased the number of judges who were both female and racial minorities as well as minorities generally. (Slotnick, Goldman and Schiavoni 2017: 395, 401). Thus, work that relies on archival data may significantly misstate current conditions.

Behavioral claims are also difficult to assess because there is not an agreed standard in the fields that study judicial behavior as to just what a truly substantive difference is. As with many elements of social science, published articles typically focus on statistically significant results. However, there is no consistent epistemology as to what counts as a robust, substantive finding. Epstein (2017: 2049) describes a 12 point gender gap in her own prior work as significant and substantial, and Rachlinski and

Wistrich (2017: 207) describe a 7 seven percent difference between judges with daughters and those without on whether to support a female plaintiff as noteworthy. Yet in the same study a 7 percent gap in the rate at which white and black defendants are convicted by white judges is described as “relatively equal” (Rachlinski and Wistrich 2017: 222).

Additionally, these claims lack context. These studies almost never compare the gap in public opinion on the same issues within the general populace. Kahan et al. (2016: 2016: 380-384, 391-397, 410-414) and Galanter (1993: 70-71) provide some comparisons between judges and the general public or other types of expert, but do not break these findings out along behavioral lines. Given that the areas which show the most promise under behavior theories are often those with broader social contexts of contestation by minority groups, this lack of comparison’s raises the question of whether judges are divided or drawn closer by their legal training and outlook. The impact of behavioral traits on judicial decision making is, at best, a model still under development.

Nonetheless, the evidence of behavioral influence is potentially directly relevant to the question of judicial decision making in the context of expert admissibility. In particular, Boyd (2013: 202-209) reports that female judges settle cases both more frequently and more rapidly than male judges do, regardless of the subject matter. This is believed to be related to patterns of socialization and management style, and the concept that female judges may be less attached to the adversarial model of dispute resolution than their male counterparts (Boyd 2013: 211-213). As controlling the admission of experts can be a means to driving settlement, gender and other behavioral traits may also

be associated with rates of rejection of proposed experts. No study in the literature has actually tried to use gender or other behavioral traits in any way in this particular context. For those reasons, the relationship between gender and judicial action in this area was investigated by assessing if there was any indication that female judges were more likely to reject expert testimony than their male counterparts as a means of facilitating case resolutions.

Another potentially relevant behavioral trait in this judicial cohort is their socialization regarding *Daubert/Kumho*. *Boyd's* examination of gender and settlement is predicated on the differentiation in socialization between male and female judges regarding the idea of adversarial processes. Similarly, Bailey and Maltzman (2011:94-100), and Songer, Sheehan and Haire (2003:125-126) both show that there are intergenerational shifts among judicial actors at the Supreme Court and the appellate court level in terms of which issues predominate and how judges use different sources of precedent. These support the idea that the time a judge is appointed, and particularly what are seen as prevailing norms at that time, appear to shape their conduct relative to prior efforts.

A number of prior studies looked at the immediate impact of *Daubert/Kumho* as doctrinal law relative to *Frye*, including Buchman (2007), Cheng and Yoon (2005), Helland and Klick (2012) and Dixon and Gill (2002). None, however, looked at whether there was any difference between judges who had been on the bench before the announcement of *Daubert/Kumho* and those who were appointed after its “gatekeeping” function was standard practice. The judicial cohort in the present study contains several

judges appointed before *Kumho* was announced who therefore experienced it as an immediate shift, and many who were socialized to see gatekeeping as a routine function. It may be that this latter cohort is more comfortable examining scientific and technical evidence in detail and therefore might be more willing to restrict or reject expert testimony.

Thus at least two different variables – gender and time of appointment to the bench – should provide some insight into the impact of behavioral traits on judicial decision making. In light of the work of Boyd (2016) expert admissibility may lack to type of identification that leads women and minorities to support the claims of plaintiffs in discrimination cases, but if nothing else that would help map the scope and significance of such influences. If these background traits have any influence, moreover, placing them alongside the legal, attitudinal and cognitive influences on judges should result in a more complete picture of how district courts carry out their daily activities.

2. The Attitudinal Model Generally and its Potential Relevance to Admissibility Determinations

By comparison to behavioral investigations, the attitudinal model has proven fairly robust in predicting judicial behavior, particularly at the Supreme Court level. First espoused by Schubert (1965), this model asserts that political attitudes drive the outcomes of disputed cases rather than such legal norms as precedent. In its simplest form, the model asserts that to the extent they can judges advocate policy views, with conservatives favoring conservative outcomes and liberals furthering liberal values (Segal and Spaeth 2002: 86-87). Embedded within the attitudinal model are concepts borrowed from economic and social theory including that actors are assumed to be

rational economic maximizers, but constrained by rules and norms (Segal and Spaeth 2002: 92-93). However, the structure of the federal judiciary and the unique position of the Supreme Court reduces the constraints of judges, and the Justices in particular, to a significant degree (Segal and Spaeth 2002: 93-97).

The attitudinal model thus takes the view that judges, and especially members of the Supreme Court, are independent actors as envisioned by Hamilton in *Federalist* 78. If anything, the model views them as robust actors in a way even Hamilton did not foresee. Where he argued their lack of control over the sword or purse would make the Court have to demonstrate sound judgment to gain adherence to their pronouncement, this model argues that such concerns do not trouble the members of the Court. Rather, they are so independent that they can operate as unconstrained policy advocates in most cases with a salient issue split (Segal and Spaeth 2002: 93-97). The only exceptions Segal and Spaeth acknowledge are when an issue poses a unique threat to that power – in which case, the members will generally side with retaining it – and making strategic decisions about whether to accept a case in the first instance based on relative power and perceived voting alignment. (Segal and Spaeth 2002: 94, 118, 168)

Similarly, the fact that removal from office and even legislative reversal is a fairly minimal threat, and that modern Justices seem to lack the political ambition of prior generations, all provide reasons the members of the Court might be less constrained than other actors. Thus, per Segal and Spaeth, maximizing policy outcomes becomes the default objective of a rational actor, at least at the Supreme Court.

At the heart of the attitudinal model is the concept of attitudes. As seen by Segal and Spaeth, attitudes are a “‘relatively enduring’ set of interrelated beliefs about an object or a situation.” (Segal and Spaeth 2002: 91). Social action results from the interaction of attitudes about an object interact with those regarding a situation. In particular Segal and Spaeth assert that cases involving similar issues build to form “issue areas” in which a judge’s conduct will be predicated on their values. Justices can be placed into a essentially linear ideological space from conservative to liberal based on their positions on these issues (Segal and Spaeth 2002: 89-90). Thus, the core premise of the attitudinal model is that how Justices vote in any given case will be a function of their pre-formed attitudes towards social groups and legal issues.

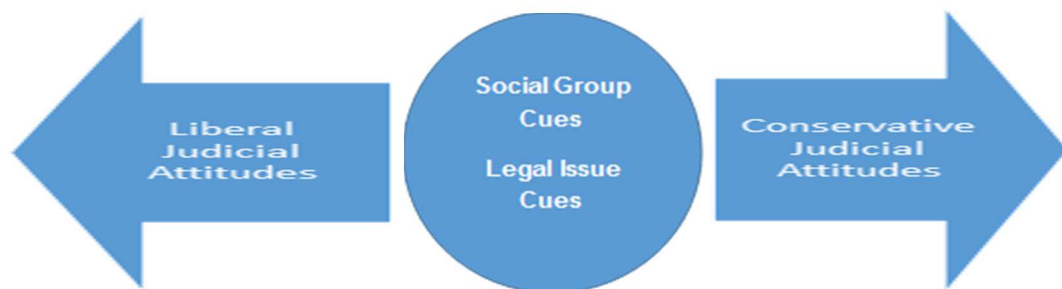


Figure 2: A Simplified Representation of the Attitudinal Model.

In a number of studies, correlations of attitudes to outcomes have been established. This seems to be a trait reflected not only at the Supreme Court but to some degree in lower courts as well (Miles and Sunstein 2008: 838-840. 846-850; Rowland and Carp 1996: 24-57). For example, judges appointed by liberal Presidents have in turn taken more liberal positions on issues like civil rights throughout the appellate court system (Songer, Sheehan and Haire 2003: 111-115). Panels composed of liberal judges vote more liberally than panels of conservative judges on issues of administrative law,

with mixed panels moderating their positions (Miles and Sunstein 2008: 846-850).

However, these results are not universal, but seem to vary by subject matter of the dispute, the circuit court that hears them and other factors.

The impact of political affiliations have also been studied with regards to trial courts. Rowland and Carp (1996: 24-57) report statistically significant if not massive variations based on whether a judge was appointed by a liberal or conservative President in a variety of areas. The magnitude of the difference is variable, with politically salient issues demonstrating a greater degree of variation. Moreover, as Rowland and Carp only use presidential party rather than more sophisticated measures of judicial ideology that account for factors such as the influence of senatorial courtesy they may understate the influence of ideology in many cases (Giles, Hettinger and Peppers 2001: 636-638).

Overall, the attitudinal model has demonstrated that in a variety of circumstances judges act in a manner that is consistent with their overall political goals. At the same time the tendency of most of these studies has been to look at issues of high political salience and conflict. Segal and Speath (2002: 324-326), for example, built their work around the Fourth Amendment specifically because it had what they saw as clean conservative/liberal linear break. Similarly, Miles and Sunstein (2008: 842-844) admit that their work on administrative law began from the premise of looking at a space with a clear liberal/conservative divide. They also admit that throughout the study of what they deem "new legal realism" the effects of political alignment, while noticeable, are often less pronounced than one might expect given the overall societal divide on similar issues.

In other words, the gap between a liberal and a conservative judge is less than that between any other liberal and conservative.

As with behavioral assessments, one question that many of the attitudinal studies fail to assess is what should count as a meaningful variance. Admittedly, any variation at all can be asserted to be at odds with the normative goals of the legal model. However, that doesn't acknowledge that variation in interpretation is part of the human condition. In other fields, experts routinely disagree with each other a significant percentage of the time (Galanter 1993: 70-71). In medicine, for example, it is common for multiple doctors looking at the same case file to disagree about whether an injury resulting from a hospital stay was a matter of negligence or not (Baker 2005). In psychology, finding two thirds or more of subjects overvalue small losses and undervalue small gains was enough for Kahneman and Tversky (1979: 265-268) to announce that prospect theory was a robust finding within the population as a whole, an assessment that has stood up over decades.

There does not appear to be any consensus in the field of judicial decision making how much variation is acceptable or publishable. As noted, Buchman (2007: 686) cites a 14 percent variation between democrats and republican judges as an indication ideology has a significant impact on judicial rulings. Setting aside the methodological questions mentioned earlier, this seems similar to many of the published findings in both behavioral and attitudinal models (Rachlinski and Wistrich 2017: 205-223). Yet Kahan, et al. (2016: 391-397) use a 75 percent congruence amongst all judicial officers as evidence that the law, rather than extralegal influences like a compelling factual narrative, drive judicial decision making. This is consistent with the congruence of lay jurors and judges used in

the original Chicago Jury Studies to vindicate the jury system (Galanter 1993: 70-71), but is certainly far from an agreed upon standard.

This ultimately raises questions as to whether establishing variations and correlations among judges reveals much about the legal system. One view is that the reliance on exceptional cases does not so much illuminate judicial decision making as unfairly criticize it against a straw man legal standard (Tamanaha 2009). The admitted lack of robust explanatory power in many of these surveys lead some legal actors to claim they are casting aspersions on the work of jurists out of ignorance (Edwards and Livermore 2009). Unless this skepticism can be surmounted, the impact of this scholarship will be limited in the very field it is studying.

This criticism has been acknowledged within the field of judicial decision making but not fully addressed (Epstein 2016: 2071-2072). Some scholars want to assert that the field is coming to recognize its own limitations and reconcile with the legal model (Geyh 2016: 51). But one would be hard-pressed to agree with that conclusion after reviewing the recent surveys of the field by Epstein (2016) or Rachlinski and Wistrich (2017). It is even more difficult to reconcile with the open hostility of Justice Roberts for sociological "gobbledy-gook" to think that the gap between political science and law will close so long as political science focuses on extreme cases and does not contextualize its results (Rocco 2017).

Despite these potential criticisms, assessing the impact of attitudes on admissibility decisions presents an opportunity to re-examine prior scholarship as well as attempt to build on it. As noted previously, only one study has suggested that ideology

matters in this context (Buchman 2007: 680-690). The operative idea in this study was that given the deferential standard of review on admissibility decisions it presents the greatest opportunity for trial judges to act in keeping with the attitudinal model. Given that conservative judges are seen as more hostile towards tort litigation generally, Buchman investigated if in the context of tort litigation conservative judges acted differently than liberal judges. Based on an analysis of decisions available through LEXIS, Buchman finds that conservative ideology is associated with a 13-14 percent reduction in admission rates relative to liberal colleagues. (Buchman 2007: 686). He also asserts that judges are more hostile to evidence proffered by individual plaintiffs (Buchman 2007: 687). At the same time, consistent with other work he finds no indication that *Daubert* substantially impacted these tendencies.

At first glance, Buchman's assertion that he conducted a national review of "all tort cases" involving issues of expert admissibility from 1983 through 2003 might seem to obviate the need for the current study, which is regional and covers fewer years. However, as discussed previously there are issues both with Buchman's methodology and results that suggest something is amiss in his data selection. Buchman's data were selected by a search in LEXIS that specified the word "tort" as one of the search parameters. (Buchman 2007: 689 n. 10). Because LEXIS's Boolean search feature only looks for the literal text, this search would only capture cases that use the term tort. However, most cases that sound in tort in Federal court do not use that term - in the data set under review in the present study, it appears in only a minority of cases. Moreover, Buchman's (2007: 681) finding of a 43.7 percent overall admission rate is at odds with

the statements that admission is to be the default rather than the exception according to both the Supreme Court and the Advisory Committee notes to FRE 702. The only study to report somewhat similar findings on this is Merlino, Springer and Sigillo (2011: 11-12, 14-15), and only as to toxicologists and engineers rather than all proposed experts. It seems likely that Buchman's data focused on toxic torts or mass tort events, areas of higher salience for conservative judicial actors and the tort reform movement (Kagan 2001: 147-150). Thus, a review of the impact of ideology and attitude on admissibility rulings that is not so circumscribed would appear to be warranted.

In general, Buchman's hypothesized relationships have facial validity. A key prerequisite of the attitudinal model is the opportunity to act in accordance with policy interests (Segal and Spaeth 2001: 93-97). Generally trial judges will assert that their ability to do so is severely limited by appellate review. (Justice [1992] 2013:46; Rachlinski and Wistrich 2017: 209-10). However, the relatively deferential treatment of expert witness admission determinations under *Joiner* provides district courts with a relatively greater freedom of action.

The other key precondition of the attitudinal model is that judges have policy preferences that their rulings will advance. Although less pronounced than the political cleavage on search and seizure doctrine studied by Segal and Spaeth (2002), there is reason to think there would be a split between liberals and conservatives on the admission of expert testimony in civil cases. Specifically *Daubert/Kumho* was initially hailed by advocates of the tort reform movement, which is generally aligned with conservative political views. (Kagan 2001: 147-150). The particular framing of the issue by

conservatives was that it would permit the elimination of “junk science” that was supposedly being used to prop up flimsy legal claims (Schwartz and Silverman 2006: 226-231). Other observers at the time and subsequently, largely aligned with liberal views of the law, have argued it created a distorted field of judicially endorsed science and artificially constrained the ability of wronged parties to bring claims based on developing scientific fields (Jasanoff 1995:130-131; Jasanoff 2008: 128-129; Beyea and Berger 2001: 348-360).

Moreover, the vast majority of *Daubert/Kumho* challenges are brought by defendants against the plaintiff’s testimony (75 percent in the current dataset, 89 percent in Buchman). Given that a plaintiff whose causation or damages expert is entirely eliminated will often have no ability to meaningfully succeed at trial, this would create a means for a judge to eliminate claims that he or she found problematic as a class even if they were permitted by law. This is in turn arguably more likely to be attractive to conservative judges in light of underlying beliefs about the proper role of law and scope of liability. This might even be further enhanced when a plaintiff is an individual and the defendant represents business or the government as hypothesized by Buchman.

Thus, the admissibility of expert witnesses presents an opportunity to assess both behavioral and attitudinal models within a setting which the outcome is unpredictable. There is not as clear of a reason to suspect socialization will play a role in the outcome as in studies with a clearer identity between the behavioral traits of the judge and the subject matter of the decision, but if it does that suggests that socialization is likely quite significant to most decisions. Similarly, if the finding of ideological influence is

replicated despite moving from Buchman's more constrained set of admissibility decisions that mention the term tort to the broader field of all such decisions it would underscore the significance of the attitudinal model. At the same time, testing these factors, the legal model, the influence of institutional constraints and cognitive shortcuts should provide a fuller sense of the relative impact of the numerous factors that can influence judicial decision making

C. Strategic/Institutional Models of Judicial Decision Making – Judges' Decisions are Constrained by Circuit Practice and Docket Pressure

Although a number of correlations have been found between judicial attitudes and verdicts, the attitudinal model has never been fully embraced. As noted, a number of defenders of the legal tradition assert the attitudinal model overstates the impact of politics and that the legal model has been misrepresented, rendering the results largely irrelevant. Other researchers, while still critiquing the legal model, have also criticized the attitudinal model is too reductionist. In particular, it has been asserted that the attitudinal model ignores the role of institutions in shaping the nature of judicial decision making (Heise 2002: 842-843; Clayton and Gillman 1999: 3-5). In this tradition, the rules and practices of a court as well as the pattern of interaction between courts and other government actors may explain deviations from what the attitudinal model predicts.

Within this grouping of institutional approaches, one that stands out as an alternative to the attitudinal model is the strategic model of Epstein and Knight (1998). This model is rooted firmly in the tradition of law and economics and assumes judges are generally acting consistently with the expectations of rational choice theory (Epstein and Knight 1998: 27-35). As with the attitudinal model the default assumption is that judges,

and specifically members of the Supreme Court, are looking to maximize their policy positions. Where the strategic account primarily breaks with the attitudinal model is in suggesting that goals other than policy might motivate actions and that even Supreme Court actors remain constrained, and as a result must act strategically to further those policy interests. For example, when a justice adopts a compromise legal standard rather than what their writings or other statements would suggest is their ideal solution, this is an example of acting strategically. Similarly, a consideration of what other political actors might do in response to their conduct is a strategic choice. Policy actions in this model are assessed against and mediated by an institutional backdrop including the criteria for setting a precedential opinion and similar norms or rules. That is, justices (and by extension all judicial actors) should seek a maximal policy advantage within the constraints of their institutional setting.

One of the most sophisticated recent analyses of the Supreme Court has identified a number of restraints that can be characterized as strategic or institutional and seem to influence outcomes in ways that vary with simple attitudinal preferences (Bailey and Maltzman 2011: 74-76, 82-87, 104-106, 138-139). One of these – deference to precedent under the doctrine of *stare decisis* – is consistent with the legal model in a broad sense. The ability of other branches to influence the Justices' behavior also suggests inter-branch relations serve as a strategic institutional constraint. In keeping with interpretivist views of institutions, the influence of these factors varies over time rather than remaining constant, with individual members of the Supreme Court generally following historic trends. This is resonant with both Dahl's (1957) conception of the

Court as a significant but limited policy actor and Keck's (2016) empirically backed assessment of the Court as a key venue for partisans to seek policy advantage and assert their rights in a way that results in more moderate resolutions of contentious issues than would arise in its absence.

Although originally focused on policy outcomes as a primary motivation of judges at all levels, Epstein and Knight (2013: 12-13) subsequently embraced the idea that the goals of judicial officers can include a number of other possibilities. As Epstein (2016: 2041-2048) notes, studies have used more than 20 different potential goals as a basis for strategic studies in addition to ideology. Of particular relevance to this study would be the personal factors of job or external satisfaction, leisure/schedule maximization and promotion. Thus, the model as a whole can be seen as a combination of a number of goals, as constrained by other actors and institutional norms:

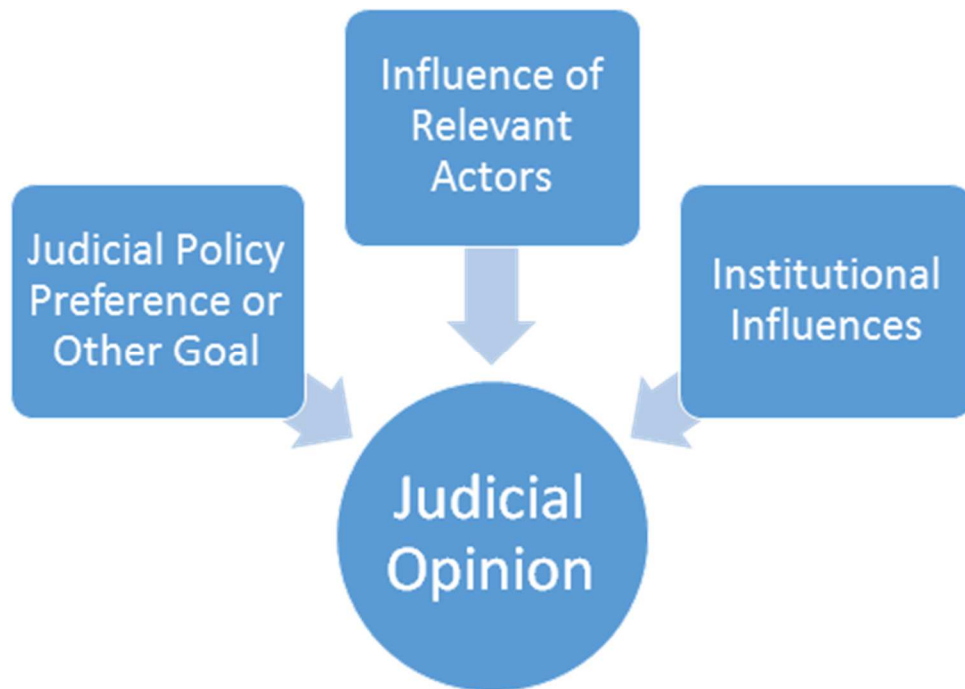


Figure 3: A Simplified Representation of the Strategic/Institutional Model.

For district court judges, both their goals and constraints are different than those that impact the Supreme Court on most issues. Policy goals can certainly matter in the context of questions like the Trump administration's travel ban or the Obama administration's actions to extend DACA style protections to the parents of "Dreamers." But these represent a tiny minority of issues on a judge's civil docket. Most are not nearly so overtly policy related, but instead may range from relatively routine questions of liability for alleged harms ranging from products liability to complicated, technically driven claims of patent or copyright infringement. Any given district court case is unlikely to trigger Congressional or executive oversight, or even much public awareness. Additionally, the uneven timing of appointments and the nature of lifetime appointments can lead to situations in which individual judges are out of step with the prevailing sentiments of their controlling circuits (Justice [1992]2013: 46-48). In that circumstance it would be irrational for a judge to act as the sort of sincere policy actor posited by the attitudinal or strategic models.

That said, one reason district court judges might be more willing to try to advance an ideological position is for the sake of "auditioning" for a promotion. Unlike the Supreme Court, lower Federal court judges have the possibility of career advancement, with district court judges being elevated to appellate courts and the appellate bench being a key feeder to the Supreme Court (Slotnick, Schiavoni and Goldman 2017: 397-399). While the odds of promotion to a higher court are low, a judge can eliminate themselves from contention if they are too idiosyncratic or act against the policy interests of a President of the same party as the one that appointed them (Posner 2008:141). This

might encourage judges to hew to norms of conduct but also incentivize some tendency to create an ideologically consistent pattern of behavior.

This calculus changes once judges take senior status. Upon taking this semi-retired position, the possibility of promotion is eliminated. At the same time, they generally surrender discretion over their case assignments to the chief judge of their district or circuit (Posner 2008: 159). At this point, a strategic model might suggest a moderation of any sort of ideological tendency as chief judgeships rotate among the active members of the bench.

Still, asserting that judges would consciously consider promotion in deciding civil cases seems unlikely if possible. A more likely goal of judges in this setting is the desire to resolve matters expeditiously. Whereas the Supreme Court can (and does) restrict the flow of cases quite effectively, lower courts lack the ability to pick most of their cases. They have to resolve the matters before them. Case filings have grown significantly over time, with ever greater resulting pressures on judges to handle matters as expeditiously as possible (Boyd 2013: 197-198; Robbennolt, MacCoun, and Darley 2010:29). Moreover, the fact that judges are generalists, and expected to deal with a wide variety of cases, also puts pressure on judges to find the path of most rapid resolution (Posner 2008: 265). Thus, consistent with the strategic model, it can be posited that judges are likely trying to do whatever they can to reduce the sheer volume of their caseloads where possible.

The primary constraints on judges' ability to do this are the expectations of lawyers and the vertical hierarchy of the Federal Court system. (Baum 2006: 50-53, 99-100; Posner 2008: 43-45; Songer, Segal and Cameron 1994: 674-677). District courts in

particular have both the Supreme Court and their own circuit precedent to contend with. Lower court judges are obligated to at least craft their decisions in a way that comports with both levels of precedent or risk dissatisfied counsel appealing. The significance of this constraint is shown by the fact district court judges overestimate the risks of reversal on appeal in most instances (Rachlinski and Wistrich 2017:10). Thus, district court judges cannot just blatantly eliminate cases without what appears to be good cause.

However, the discretion given to judges over pre-trial rulings on everything from evidence to dispositive motions creates a potential means for them to further the goal of reducing their docket pressure consistent with this constraint. (Rowland, Traficanti, and Vernon 2010: 185-189). As Kim, Schlanger, Boyd and Martin (2009) note, the time demands any particular matter places on a judge varies depending on the docket context of the matter. Many cases are resolved through party settlement discussions with little or no court intervention. At various phases in the life of a case, though, more burdens are placed on a court, ranging from hearings to writing orders to conducting trials. At some points, such as motions to dismiss or motions for summary judgment, the parties offer the judge an opportunity to avoid more costly future burdens such as a trial by resolving the matter through a dispositive ruling. Thus, a strategic account of district court decision making should consider if the timing of a particular decision provides a way for a court to further their goals of efficient case disposition and assess if the judge acts to further that goal.

At the summary judgment stage, for example, a judge is given an incentive to find a reason to eliminate the case. If the motion is granted, the matter is taken off the judge's

docket without the burden of scheduling and conducting a trial. The grant of a summary judgment motion will generally be reviewed *de novo* on any appeal, which might elevate the perceived risk of taking this action to some degree (Rachlinski and Wistrich 2017: 209-10). At the same time, that reversal is months if not years down the road, the parties will have more time to settle, and a negative result at most puts the court in the position they were in before the grant of summary judgment.

By comparison, if a case is at a pre-trial phase, rulings on motions *in limine* are highly unlikely to dissuade the parties from proceeding. At most the judge can shape the scope of the trial, which can have a significant impact on the outcome (Rowland, Traficanti, and Vernon 2010: 185-189). However, it does not come with the same degree of burden reduction. Moreover, although review of trial outcomes and fact based rulings is typically deferential, the cost of a significant error in pre-trial rulings can be much higher than at the summary judgment stage, including re-trying the entire matter. In that setting the judge may be incentivized consistent with strategic modeling to take the path of least resistance in all but the clearest cases justifying intervention.

The relationship of docket stage to judicial decision making has not been robustly studied. Most studies of judicial decision making at this level look at outputs, either generally, such as Beebe's (2006) study of trademark decisions, or at a specific moment in the docket, such as Boyd's (2013) study of summary judgment decisions in discrimination cases. There have not been many studies to compare how judges treat the same or similar issues at different moments in the timeline of a case.

This holds true for studies examining the decision making in the context of expert

witness testimony. Although Dixon and Gill (2002:294-296) suggested that co-presenting a motion for summary judgment with a motion to exclude testimony could have a meaningful impact on case outcomes they did not test it empirically. Merlino, Springer and Sigillo (2011) did not differentiate between dispositive motions and pre-trial rulings. Buchman (2007) specifically included all types of rulings in his analysis. Helland and Klick (2012) and Cheng and Yoon (2005) both had research designs that could not examine this factor.

This oversight is unfortunate as *Daubert/Kumho* rulings create a unique set of incentives at the summary judgement stage. Under *Joiner*, regardless of timing the decision to reject or admit expert testimony is given a deferential abuse of discretion review. If a summary judgment motion can be resolved based on the absence of admissible expert testimony, this would effectively reduce the risk of reversal on appeal to what would obtain if they judge let the case proceed to trial while eliminating the time and effort of a trial. Thus, a rational maximizer should be more likely to grant motions to exclude when they are made in the context of a summary judgment proceeding.

Another overlooked strategic dimension of these rulings is circuit of origin. Songer, Sheehan and Haire (2003: 125-126) identified that most differentiation in Federal law is circuit based rather than regional. However, no empirical study of *Daubert Kumho* rulings has compared decisions by judges in different circuits. Merlino, Springer and Sigillo (2011) compile cases nationally, without initially attempting to identify if there are circuit differences of note. Dixon and Gill (2002: 281) coded circuit locations but did not actually compare outcomes based on circuit of origin, only reporting a single

exemplary model for the Third Circuit. Helland and Klick (2012) and Cheng and Yoon (2005) compared state and Federal practices in discrete areas but not cross-circuit differences. Buchman (2007:147-150) looked for and did not find strategic conformity by district judges whose ideology differed from the mean value of their reviewing circuit, but did not investigate if there were any doctrinal legal variations that would indicate one circuit or another might be providing different incentives and constraints for judicial behavior.

These oversights are especially problematic in light of the fact that there are some circuit doctrinal splits that could lead a strategic actor to alter their behavior. In particular in the two decades since *Kumho* the Ninth Circuit has adopted some unique positions that could influence lower court behavior. For example, a failure to make adequate findings supporting exclusion or admission of expert testimony is grounds for a new trial, whereas in other circuits it leads to a limited re-examination procedure on remand (*Mukhtar v. California State University, Hayward*, (9th Cir. 2003), 299 F.3d 1053: 1064-1066). Similarly, the Ninth Circuit overtly permits the use of non-testifying “advisors” by judges as a basis for admissibility decisions where others do not (Cwik, et al 2013: 337-338). Multiple articles have also suggested that this circuit adopted an unduly harsh standard on causation testimony in toxic torts, which might have led to a higher rate of rejection of experts in such fields as toxicology in the Ninth Circuit. (Beyea and Berger 2001: 349-350; Beecher-Monas 2011: 42-43). All of these features combine to suggest that despite the deferential standard of review granted these rulings under *Joiner* generally, judges under the Ninth Circuit’s review might be more inclined to rigorously review proposed

expert witnesses and act to strike or restrict testimony based on the perceived risks of failing to do so. This would be consistent, for example, with the idea that the Ninth circuit as the principal permits less “shirking” with regards to the gatekeeping function and therefore a rational judge would apply close scrutiny to experts (Songer, Segal and Cameron 1994: 674-675).

Finally, although it seems unlikely that promotion decisions might hinge on how a judge handles this type of motion, there remains the possibility that the elimination of that possibility might influence their conduct. That is, relieved of even the potential for promotion might reduce the incentives to act in a manner consistent with partisan interests. Thus judge on senior status might be expected to act more similarly and with less reflection of a partisan gap.

Thus, the question of expert witness admissibility presents a number of avenues for investigating the influence of strategic and institutional considerations. The influence of when a challenge is made to an expert, what court it arises in and whether the judge has any potential for promotion all may help map the judicial decision making process. In particular placing these influences in context with a study of legal, behavioral, attitudinal and cognitive influences should lead to a more holistic understanding of when or if institutional and strategic considerations are relied on by district courts.

D. A Cognitive Model of Judicial Decision Making – Judges’ Decisions are a Function of Heuristic Processes

As discussed above, different models of judicial decision making vary in the influences they consider. However, they can all be argued to assume the judges consider all of the relevant information before them, even if they are ultimately led to different

conclusions based on behavioral traits, attitudes or institutional factors. The legal model, as espoused by Wechsler (1959) and Dworkin (1977, 1988) in particular, assumes a judge or panel is operating with comprehensive knowledge in an effort to derive lasting, neutral principles of law. The attitudinal model as described by Segal and Spaeth (2002) is predicated on the idea that judges are policy maximizers. Similarly Epstein and Knight's (1998) original conception of the strategic model was derived from classical rational choice principles of policy or other goal maximization coupled to conscious awareness of and consideration for constraints such as other political actors.

As Simon (1955: 101-105) suggests, this sort of comprehensive rationality is unrealistic in most settings, rather than rationing of limited cognitive resources. Shapiro (1965: 136-138, 14-154) argues that this trait, as reflected in March and Lindblom's conceptions of incremental decision making, provides a better fit for understanding the conduct of appellate courts than either legal process or judicial attitude, as both assumed something closer to comprehensive rationality. He also asserts that Llewellyn's theories of legal realism are in fact essentially incrementalism and bounded rationality by another name. In the intervening decades, cognitive psychology has established that in fact people tend to have two distinct systems for decision making – one that makes rapid, intuitive judgments and a second that considers matter systematically (Kahneman 2011: 12, 21-22; Baum 2010: 17). Moreover, people are generally predisposed to use less, rather than more, of their limited mental capacity to accomplish cognitive tasks (Kahneman 2011: 35). However, Shapiro's views have largely been overlooked with the rise of alternative models of judicial decision making that likely seemed less complicated

than incrementalism's process of sequential limited relative judgements.

However, it can be argued that much of the legal system is built around a concept of heuristic simplification that favors incremental, reflexive action. *Stare decisis* is a principle that is typically justified by how it impacts the expectations of society, but it also permits judges to minimize their own cognitive efforts in cases involving "settled law" (Carter and Burke 2010). While judges sometimes report that they are less constrained by precedent than lawyers are prone to think, the hierarchical nature of the court system and precedent does help define and generally limit the universe of choices on which to base an argument (Posner 2008: 43). Compared with the open-ended nature of a literature review for a dissertation, especially an interdisciplinary one, the confines of the several hundred volumes of the United States Reporter or even the several thousand that make up the Federal Reporter seem almost cozy by comparison.

Appellate judges also have the cognitive simplification of an agreed upon universe of facts – the record and appendix on appeal (Carter and Burke 2010: 10-12, 28-29). Technically the record below contains all materials considered by the court below, including transcripts of proceedings, rulings on motions and the evidence itself. However, by rule in most appellate courts the parties designate only the aspects of the record necessary to support their positions. Where allowed, the parties will place the most critical material as an appendix to their principal brief or a joint filing to make it even easier for the court to limit its effort. While in theory judges can review the entire record themselves, and can seek outside sources of information, this is the exception and not the rule.

Trial courts lack the luxury of a similarly circumscribed universe. However, they have a set of tools that can constrain the actions of both parties and judges, helping to reduce cognitive burdens to tolerable levels (Posner 2008: 176). For example, the Federal Rules of Civil Procedure, local rules and scheduling orders to lay out the course and scope of a case and limit the types of decisions they have to make in each stage. The rules of evidence similarly make clear what is generally acceptable, and parties are encouraged to agree to the scope of testimony and number of witnesses to narrow issues to those that are truly contested. Mandatory disclosures, exchanges of discovery, expert reports, case management conferences, summary judgment and other pre-trial motions are also used to identify and simplify trial issues.

In addition to circumscribing the world of relevant material generally, many of these procedures also create what can be seen as decision heuristics. For example, a motion for summary judgment under Federal Rule of Civil Procedure 56 is one of the most common dispositive motions – so-called because a ruling at this stage can dispose of a matter in part or in whole. The general standard for summary judgment is deciding if there is any issue to be resolved by the jury, which on its face would seem a daunting task. However, it is tremendously simplified by the rule and precedent associated with it - the judge must assess the evidence in the most favorable light to the non-moving party, can only grant the motion if there are no material questions of fact to be resolved and cannot base the ruling on their own interpretation of the weight or credibility of admissible evidence (Fed. R. Civ. P 56). This is frequently supplemented with local district or court rules that speak to the format, length and nature of materials to be

considered and require the parties to set out for the court both what facts they consider uncontested and which areas, if any, they assert create material questions of contested fact that would defeat the motion.

While the sheer length of most rulings on summary judgment motions speaks to the fact that these procedures can still leave a judge with many issues to consider, they are operating within a constrained space. Most judges have a template for their rulings, frequently with the law associated with the procedure already filled in, subject to supplementation with a discussion of the contested substantive claims and with most of the body of the ruling focused on a handful of key assessments. These are typically assessed on an element by element basis, with the judge breaking setting out the position of the parties and then providing his or her own assessment as to whether there is a basis for granting the motion. All of this serves both to clarify the basis for the action and to simplify the cognitive task.

Another area of cognitive simplification in the law involves simplifying complex legal formulas. In many areas of law, such as preliminary injunctions, trademarks, and patent damages, the Supreme Court or an appellate court will announce a multi-factored set of considerations that can be applied to resolve a matter. Sometimes, the controlling law states all the factors as have to be considered, such as the requirement in a preliminary injunction context that the judge assess the likelihood of success on the merits, the likelihood of irreparable harm, the balance of harms between the parties and whether the grant would be consistent with public policy and the interests of justice.

Winter v. Natural Resources Defense Council (2008), 555 US 7. Different permutations

of the same standard can vary the outcome- some circuits allow for any success on the merits to be enough, whereas others place a higher burden on a moving party – but they consistently state that all factors have to be considered. In these circumstances judges go through each factor in order, and treat them as an overall balancing act.

By comparison, in trademark law, while there is an agreed upon standard for infringement – likelihood of confusion between a senior and junior use of a mark in commerce – there is a wide variety of judicial implementations of that standard in different circuits. As Beebe (2006: 1581-1582) notes, the number and type of factors to be considered varies by circuit, and none of them sets forth an exhaustive list of factors to be considered. Instead, case law typically sets out a series of considerations that can be used in an appropriate dispute, while asserting none are dispositive. In other words, a set of potentially sufficient conditions ranging from the visual similarity of the marks, to whether there is evidence of conscious copying, to survey evidence about the impression of an “ordinary consumer” when seeing the marks in context. In practice, however, Beebe reports certain factors – the absence of similarity or a lack of direct competition – are typically dispositive in a defendant’s favor, while bad faith is generally dispositive in favor of the plaintiff. Other factors are largely irrelevant, and what is perceived as the gold standard of evidence – consumer surveys – are almost never used.²

A similar pattern, albeit with a different origin, has been identified in the context of securities law. There, judges have developed a set of maxims that tend to serve as the

² Although it was not a specific focus of the current study, there is evidence in the cases reviewed that surveys are an area of expertise in which disputes over methodology are frequent. Given Beebe’s findings it is unsurprising that parties would avoid such costly effort except in unusual or critical instances.

basis for dismissing claims (Bainbridge and Gulati 2002: 85-88). Based on the highly technical nature of securities law, the general lack of familiarity with it that judges possess, and the motivation to resolve matters promptly to reduce docket pressure, Bainbridge and Gulati postulate these doctrines serve as substitutes for a more holistic assessment of claims. This assertion is not robustly documented empirically, as the authors primarily rely on assessing and comparing the plausibility of the heuristic impulse with an alternative, Dworkian rationale for the individual doctrines rather than documenting the frequency with which these doctrines are invoked, but it does have some facial validity.

While simplifications are a rational means of simplifying complex tasks, some simplifications can result in notably sub-optimal outcomes. These findings are at the heart of much of the work in behavioral economics by authors like Daniel Kahneman, Amos Tversky and Daniel Ariely. One team of researchers – Chris Guthrie, Jeffrey Rachlinski and Andrew Wistrich - have led the effort to assess how these previously documented issues may influence judges. Taking advantage of unique access to judicial officers (based in part on the fact that Judge Wistrich has served as a United States Magistrate judge for a number of years), they have repeatedly been able to demonstrate that judges are prone to a range of heuristics (Rachlinski and Wistrich 2017: 211-216). They have also shown that judges, in some settings, are less prone to heuristic behaviors of the general public, leading to questions of when and how judicial conduct can be differentiated.

One of the earliest reported findings of Guthrie, Rachlinski and Wistrich (2007) was that judges were, at least in experimental settings, similar to the general public and members of other professions. In one noteworthy example, judges performed on the Cognitive Reflection Test at the same rate as many other professions and educated people. The CRT consists of three seemingly simple questions that call for logic to avoid falling into intuitive traps. For example, it asks “If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets?” (Guthrie, Rachlinski and Wistrich 2007: 10). The correct answer is 5 minutes, but the symmetrical, intuitive answer invited by the question is 100 minutes. Only a small minority of the judges surveyed got all three answers right, and in fact the mean score on the 3 questions was 1.23 correct answers. This is well below MIT students, but comparable to most other professions and schools for which data exists (Guthrie, Rachlinski and Wistrich 2007: 13-16; Rachlinski and Wistrich 2017: 211-212).

Guthrie, Rachlinski and Wistrich’s work has also shown that these intuitive, relational impacts extend to tasks specific to the judiciary (Rachlinski and Wistrich 2017: 212-216). Anchoring, for example, is a well-documented heuristic in which the introduction of even arbitrary numbers can have an impact on how people interact and establish values under conditions of uncertainty. For judges, the introduction of different random settlement demands in a hypothetical bench trial – demands which are supposed to be irrelevant to the final assessment of damages – nonetheless had a material and predictable impact on the final number. Where the initial settlement demand was low, the typical plaintiff received a fraction of what the high demand plaintiff obtained

(Wistrich, Guthrie and Rachlinski 2005: 1286-1293). At the same time, the same experiment showed judges were generally better at ignoring some types of inadmissible testimony in reaching a result, whereas jurors tend to be influenced by the reason the evidence was excluded, its salience and content. (Wistrich, Guthrie and Rachlinski 2005: 1320-1322). Thus, there is something unique about judicial behavior, likely related to training and experience.

Of particular relevance to the current study, in another experiment the same team found that judges assess experts relatively rather than individually consistent with contrast effects. A contrast effects can lead to overvaluing an object or person based on the presence of a poor alternative, and is often used in sales to convince customers an extra service they may not want is really a bargain because of the seemingly low relative price. In the context of experts, Rachlinski, Guthrie and Wistrich (2013: 1597-1601) reported that inserting a weakly credentialed expert's biography into a packet of material led to much more positive assessments of other experts' testimony. Essentially, the "bad" expert is not only disregarded, their better regarded peers are more likely to be believed. This effect held not just for opposing experts but all experts in the scenario, suggesting it might be worthwhile for a party to include a range of experts and use them strategically.

Ultimately, this line of work both provides some useful context lacking from much of the other models and suggests that applying the teachings of cognitive science to judicial opinions is likely to be a worthwhile endeavor. Unlike the other models surveyed, the cognitive work starts from the premise that judges are, first and foremost,

humans – a finding that overtly echoes the work of Jerome Frank (Rachlinski and Wistrich 2017: 223). Thus, while in some settings they may be legalists, policy actors, or strategic thinkers, they are just as apt to simply be attempting to navigate uncertain decisions as well as they can. These findings also underscore that judicial behavior is generally similar to that of the general public, and where it is not it tends to fall in line with judicial training and norms. This at least suggests that some of the broader assertions of the attitudinal model about the centrality of policy considerations should be tempered, and that where possible researchers should bear in mind the possibility that what is being observed as significant variations are no more than normal human variations in approaches and perspectives.

Although Guthrie, Rachlinski and Wistrich have conducted a number of experiments with significant internal validity about how judges approach decision making generally, this work has not been conducted in the native setting of judicial actors. Judges were asked to engage in both standard cognitive tasks such as the Cognitive Reflection Test and to offer judgments on hypothetical scenarios, but these experiments were conducted at conferences and through surveys rather than in the actual context of cases. This specifically removed one key element for much decision making under conditions of uncertainty – risk. As Kahneman and Tversky (1979) documented in prospect theory, the relative and asymmetric weighting of risk typically leads people to be averse to loss to a greater degree than they are risk embracing for the possibility of similarly sized gains. This pattern reverses itself in cases of extremely large and unlikely

losses or gains, with participants becoming risk seeking for the possibility of avoiding a certain loss.

This may not seem relevant at first glance to judicial decision making, as judges face no financial consequences from their conduct. However, reversal on appeal is a form of risk that judges cannot avoid. A reversal not only carries professional stigma, as discussed above it typically means more work for the judge. In the context of admissibility of experts, this might range from having to craft new findings to justify their actions all the way through overturning an entire trial under the Ninth Circuit's *Mukhtar* standards. Thus, risk is existent in judicial decision making in the real world, and assessing if judicial behavior is consistent with human conduct generally under such circumstances should be explored.

To date relatively little work has sought to explore the impact of heuristics on judicial rulings. Baum (2010: 17) hypothesizes that heuristics are at work in asserting that judges are likely subject to a variety of motivations that should be studied. Epstein and Knight (2013: 12-13) also suggest a role for heuristics as part of a proposed motivational approach. Neither provides an analysis of the operation of heuristics in any particular setting. Beecher-Monas (2001: 1589-1590) proposes a checklist for courts to use as a heuristic in the context of *Daubert* hearings, but provides no analysis of actual judicial conduct in these hearings. No study seems to have looked at the actual conduct of judges in the context of admissibility decisions with a focus on the whether judges engage in heuristic processes or what those processes might be.

Given the findings of Guthrie, Rachlinski and Wistrich, decades of research in general cognitive psychology and the series of heuristic simplifications embedded in the legal process, it would be surprising if courts did not adopt a heuristic process in addressing admissibility questions. There is some evidence suggesting that a heuristic model explains the process as well or better than other models. Gatowski, et al (2001: 444-448) report that judges understand certain *Daubert* factors – general acceptance and peer review – much better than they do more complex issues such as replicability and error rate. Dixon and Gill (2002: 285-288, 299-300) find that general acceptance is the *Daubert* factor judges are most likely to analyze, and that it was predictive of judicial conduct in most cases. This is also consistent with Merlino, Springer and Sigillo's (2011: 17) finding that general acceptance is a key factor in the determination of admissibility, and with Helland and Klick's (2012: 32-33) finding that courts expressly using general acceptance as the controlling standard have similar numbers of testifying experts as courts applying the flexible multi-factored standard of *Daubert*. This suggests that judges likely consider how an expert's testimony fits within their field as a threshold or benchmark consideration before assessing whether to investigate other, more complex *Daubert* factors. This would also be consistent with jury studies showing jurors tend to fall back on "peripheral processing" issues like signals of trustworthiness or sincerity, rather than the scientific content of expert testimony, particularly in complex cases (Cooper and Neuhaus 2000: 168-171). For judges, general acceptance may well replace cues like body language, dress or perceptions of the speech or attitude of a witness relied

on by jurors. However, expert witness admissibility has not been explored with a distinct focus on whether judges are adopting a heuristic.

Similarly, a heuristic approach may explain the differentiated rates of admissibility linked to the type of evidence noted by Merlino, Springer and Sigillo (2011: 9-15). This study looked at four categories of evidence – toxicology, damages, engineering and psychological/psychiatric. The study found that while in the other three settings a majority of experts were admitted (although damages experts were lower than the other categories) toxic tort causation evidence was rejected in the overwhelming majority of the time – in 78 of 100 cases. This may well have its roots in a heuristic, derived not from *Daubert* itself but issues of relevance considered on remand which permitted judges to avoid tricky questions about scientific methods in favor of simple counting.

In the *Daubert* case, the central question was the standard of admissibility for toxicology reports based on epidemiological evidence relating to anti-nausea medication during pregnancy and its potential to cause birth defects. The plaintiffs relied on the toxicology reports and the fact of the obvious birth defects to support both general (that is, the possibility that the drugs could cause harm generally) and specific (that the harm was caused in this particular case) causation. On remand in *Daubert*, the Ninth Circuit was not only focused on scientific methods but the relevance of the evidence to the case as a whole (Beyea and Berger 2001: 348-350; Beecher-Monas 2011: 41-43). Relying on California's "more likely than not" standard of causation, the panel ruled that the reports could only suffice to establish causation if they demonstrated at least a doubling of risk.

Short of such a showing, the reports were potentially evidence of general causation but ultimately insufficient to support specific causation.

Nominally, this finding was linked to a particular state common law standard of proof and theoretically about relevance rather than the general admissibility of such evidence as matter of methodology. However, per Beyea and Berger (2001) and Beecher-Monas (2011), other courts seemed to focus on the numerical standard to the exclusion of anything else. This potentially created a simple heuristic that would both eliminate the overwhelming number of studies and allow courts to sidestep nuanced analysis of evidence. As will be discussed in this study, this standard did not eliminate pharmacology related product liability claims, but it may have led to a fundamental shift in what sort of evidence plaintiffs use to make their claims and minimized the use of toxicology reports despite their potential value.

Another heuristic for judges related to admitting experts may lie in deciding whether they need to intervene at all. Cecil and Willigang (1993: 5, 20-21) report that judges are reluctant to be seen as intervening in the adversarial process through measures such as the appointment of expert witnesses. Both *Daubert* and the Advisory Committee notes to Rule 702 stress the availability of actions including cross-examination, presentation of countervailing witnesses and jury instructions as means to address weak expert witness evidence without excluding it. The Advisory Committee notes also assert admission should be the rule, rather than the exception. This may well bias judges in favor of avoiding taking a hard look at complex questions of methodology, so long as they believe the procedural fairness of the adversarial process is not threatened. Thus,

there is a strong possibility that analyzing judicial opinions regarding the admissibility of expert witnesses both qualitatively and quantitatively will support a finding that judges are using multiple approaches that minimize the cognitive effort required to assess these potentially complex questions.

One complicating factor in assessing the role of heuristics in judicial decision making may lie in the fact that any indication of the judge's reasoning is likely to be in a written opinion. As Rachlinski and Wistrich (2017: 223) note, in keeping with the work of Oldfather (2008), the process of writing can allow for greater reflection and potential for de-biasing. It can also potentially obscure the basis for the initial judicial hunch, and instead reflect a post-hoc rationalization that minimizes the evidence of the operation of heuristics. The heuristic process could be further obscured if the judge delegates writing of the opinion to clerks, who are thought to be more legalistic and formal in their approach given their lack of real world experience (Posner 2008: 221; but see Bainbridge and Gulati 2002: 86, 103 (asserting clerks have incentives to use heuristics)). At the same time, this suggests that if analysis of these opinions still indicates heuristic processes are being employed, then the heuristic model is a good fit with this type of judicial decision making. Overall, analyzing opinions for heuristics should not only provide a greater depth of understanding than studies that look solely at outcomes but also illuminate how and why decisions by district courts on this issue are either consistent with or vary from the expectations of the legal, behavioral, attitudinal and strategic models of legal decision making.

E. The Research Questions about Expert Witness Admissibility this Dissertation Explores and the General Research Plan

In light of the state of the literature, several research questions emerge around the common objective of improving understanding of judicial decision making in the Federal district courts. In particular, the following research questions frame the data selection and investigation using both a qualitative and quantitative lens:

1. Do judges as a group follow the *Daubert/Kumho* line of precedent as well as Federal Rule of Evidence 702 and admit experts as a rule – that is, do judges regardless of location, ideology, gender or other personal or institutional factors tend to follow the law in the first instance?
2. Do district courts in different circuits demonstrate different patterns of behavior regarding the admission or treatment of expert witnesses in civil cases when factors such as ideology, gender and other personal or case factors are controlled for, and if so do these run counter to a general tendency to admit expert witnesses?
3. Are there conditions under which the political attitudes of judges are correlated with variations in their decisions about the admissibility and scope of expert witness testimony in civil cases?
4. Is there a relationship between the gender of the judge and likelihood of rejection of an expert witness, either overall or in the context of dispositive motions?
5. Is there an identifiable relationship between the procedural context of a challenge to the admissibility of expert witnesses in a civil case and its outcome? Specifically, is the likelihood of success in challenging an expert greater at the summary judgement stage of a case relative to pre-trial proceedings?
6. Do judicial opinions on the determination of admissibility of expert witness testimony suggest the operation of heuristics in these determinations, such as a bias in favor of admitting experts who are within the mainstream of their field?

1. The Overall Research Plan

In keeping with the objectives and research questions posited above, this dissertation presents a non-experimental, cross-sectional comparative research study analyzing Federal trial court determinations of the admissibility of expert witnesses in

civil cases. The study relies on thematic analysis of judicial rulings on these issues from nine United States District Courts over a six year period (rulings issued from 2010 through 2015), analysis of the broader context of the motions, and publicly available demographic information about the judges under study. Both qualitative and quantitative analysis of the resulting data is employed sequentially.

This study sought to overcome a number of methodological gaps and questions raised by prior studies regarding judicial decision making generally and the treatment of expert witnesses in particular. Initially, by looking not just to outcomes but the content of judicial decisions, the study sought to get a fuller understanding of the mental processes of judges than could be achieved by solely investigating the outcome as it related to judicial characteristics such as gender or politics. At the same time, it avoided the risks of faulty or biased recall associated with interviews and should serve as a conservative indicator of heuristic tendencies. Additionally, by capturing all of the rulings within a constrained set of districts, the study sought to overcome the potential for distortion from assuming uniformity across different geographic jurisdictions. Employing qualitative thematic analysis as a first step permitted the author to identify latent issues in the data that had been overlooked by prior research, such as the tendency of courts to only rarely actually use the specific *Daubert* factors in lieu of simplified standards and the tendency of judges to restrict the testimony of experts rather than exclude them completely. These discoveries were then incorporated to refine a set of hypotheses initially derived from prior articles and used as the basis of a quantitative

assessment of how well the dominant models of judicial decision making explain this area of judicial decision making.

2. The Courts and Cases Assessed – Preserving Diversity of Appellate Oversight and Similarity of Overall Cases

One of the critiques leveled by legal scholars is that political scientists give inadequate attention to the content and context of legal rulings. (Tamanaha 2009). This has held true in studying *Daubert*, especially in the context of the attitudinal model. For example, Buchman (2007) looks solely at the outcome of cases. At the same time, studies that have actually examined the content of legal rulings have not explicitly considered the influence of location in favor of limited sampling such as Merlino, Springer and Sigillo (2011). This approach has the impact of assuming the existence of a common legal model and conditioning throughout the country as a whole, rather than bearing in mind the potential for circuit based differences in outcome. As noted above, these differences can give rise to behavior that is both consistent with the legal model and lead to strategic evaluations by district court judges depending on the nature of the difference. Dixon and Gill (2002: 281) use circuit controls and assert that they could have revealed different patterns in different circuits, but do not report those distinctions.

Thus, one of the goals in designing this study was to ensure that it captured the potential distinctions at the circuit level in a meaningful manner; preserving this potential diversity of practice while ensuring a relatively balanced sample supported a purposive sampling approach to case selection. (Seawright and Gerring 2008: 300-301).

Additionally, there was a concern that inclusion of certain atypical districts – such as the Northern District of Illinois, which is dominated by Chicago and dwarfs other districts in

the Seventh Circuit, or the Eastern District of Texas, which has become a center for patent filings to the exclusion of much of the rest of the country – would fundamentally skew any data. Instead, nine district courts - 3 geographically adjacent districts in each of the Ninth, Seventh and Fourth United States Courts of Appeals – were selected. In each case, the cluster included a mixture of urban and rural areas, so as to preserve a broad cross section of potential litigation. They included the District of Oregon, Western District of Washington and Eastern district of Washington within the Ninth Circuit, the Northern District of Indiana, Southern District of Indiana and Southern District of Illinois in the Seventh Circuit, and the Eastern, Middle and Western District of North Carolina in the Fourth Circuit.

The inclusion of the District of Oregon and the two Washington districts was predicated in part on the fact that the Ninth Circuit is perceived as operating in a distinct manner on this general subject matter. As noted above, it is the origin of potentially more stringent guidance on the admissibility of causation evidence. (Beyea and Berger 2001: 348-350). It has also endorsed the concept of a judge obtaining independent advice from non-testifying “technical advisors,” which in theory could result in judicial officers being more capable of informed critique of experts. (Cwik, et al 2013: 337-338). Additionally, it provides a harsher penalty for district court procedural errors in analyzing these questions, requiring a new trial as the result of a faultily reasoned exclusion or admission. (*Mukhtar v. California State University, Hayward*, 299 F.3d at 1064-1066). Therefore, comparison of the behavior of judges who are subject to 9th Circuit review to other districts in circuits lacking such unique patterns of behavior should help identify if

there are any meaningful trends or if judges are generally the same in their behavior regardless of location.

Identification of the other two clusters was based in part on a desire to gain a broader range of potential appellate influences while creating similar caseload profiles. This led to identifying the Southern District of Illinois, Southern District of Indiana and Northern District of Indiana from the perceived “moderate” Seventh Circuit. Unlike the Ninth Circuit, the Seventh is not known for a unique approach to evidentiary issues. Similarly, while the Fourth Circuit has historically been seen as a “conservative” court, it was not identified initially as having a unique treatment of expert witnesses.

While presenting unique potential appellate influences, the research was designed to try and avoid geographic or case type dissimilarities. As noted above, courts with a unique, dominant type of case were consciously excluded. Each cluster also provided a similar geographic cross section. Each contains one district dominated by major metropolitan region – Seattle in the Western District of Washington, Indianapolis in the Southern District of Indiana and Charlotte in the Middle District of North Carolina. Each contains a district that includes one significant but secondary city with a large population and is otherwise more scarcely populated - Portland in the District of Oregon, Fort Wayne in the Northern District of Indiana and Raleigh in the Eastern District of North Carolina. Finally the remaining district in each cluster (the Eastern District of Washington, Southern District of Illinois and Western District of North Carolina) lacks any large cities although each has some regional population centers such as Spokane, Washington; Springfield, Illinois; and Asheville, North Carolina.

Given the focus of this study on the typical conditions under which rulings are made, the goal was to identify every written opinion issued from 2010-2015 that involved the admission of expert testimony in a civil case within each district under study. To attempt to ensure all possible opinions were included, initially, the PACER (Public Access to Court Electronic Records) system was employed to identify opinions issued from the District of Oregon during 2015 on a month by month basis. This system lacks a key word search or similar function, so the cases embracing a *Daubert* issue were identified by manual review. The cases so identified were then compared with the results of a LEXIS search for the District of Oregon for the same time period that specifically mentioned “*Daubert*” or “Rule 702.” There was complete capture of the former set in the latter search, creating confidence that LEXIS could be used to compile the datasets for this project.

Three datasets were compiled for initial qualitative analysis. A set of LEXIS searches designed to maximize the capture of potentially relevant cases were conducted. In each case the search was for the same date range of 2010 to 2015 and limited to cases referencing the districts of interest. This time the initial search was for all cases mentioning the terms *Daubert*, *Kumho* or the number 702 within 25 words of the term expert. This was in turn narrowed to cases with one of the terms object, exclude, strike or limit within 10 words of the term expert. Cases were then reviewed to exclude all non-district court rulings (including both appellate court and bankruptcy court proceedings), criminal cases, those that were from irrelevant districts, were duplicative or lacked any substantive analysis of the question of expert admissibility (such as scheduling orders,

instances of summary adoption of a magistrate's recommendation or cases in which the ruling was held in abeyance pending trial).

The cases were originally coded individually within atlas.ti. The focus was on the content, context and outcome of the cases at this stage. Thematic analysis was employed to review and assess the material. The results of this analysis are reported in Chapter 3.

For the quantitative analysis, it was decided to code the outcomes at the ruling level rather than the case level. This is consistent with Buchman (2007) and allows for normalization between judges who resolved each challenge to a particular expert in a separate opinion and those who preferred to issue omnibus opinions covering a number of motions. It was also decided at this stage to exclude the only case in which an appellate judge sat by designation.

This process yielded a total of 292 total cases. 122 originated in the Ninth Circuit, 120 in the Seventh Circuit and 50 in the Fourth Circuit. When coded at the individual expert level there were 510 total rulings – 212 from the Ninth Circuit, 206 from the Seventh and 92 from the Fourth Circuit. Of these, it was ultimately established that 81 were rendered by magistrate judges while the remaining 429 were made by Article III appointees. These were then subjected to statistical analysis as described in Chapter 4.

As described in greater detail in the chapters that followed, both qualitative and quantitative analyses provided insights into the judicial decision-making process. Initially, thematic analysis suggested that judges do generally follow the expectations of the *Daubert/Kumho* line of cases that admission of expert witnesses should be the default.

It also suggested that judges are prone to a variety of cognitive shortcuts, but do attempt to police extreme instances of witness or attorney behavior. This in turn refined a set of hypotheses that ultimately suggest that legal principles are the dominant but not exclusive influence on judicial decision-making in this context. Ideology matters, but so does circuit of origin and the nature of the challenge levied against the expert. Although it could not be established conclusively with this data set, there is also a suggestion that when a challenge is brought might impact the likelihood of an expert being rejected. Ultimately, no one model fit the evidence perfectly, but several provide complementary pieces of the mosaic of district court decision making in this context.

Ch. 3 - Thematic Analysis Suggests Judges Favor Legal Norms Over Strict Application of *Daubert/Kumho*

The qualitative analysis in this case followed the precepts of thematic analysis. (Braun and Clarke 2006). Beginning from the framework of directed content analysis, initial codes were derived from prior theory and empirical work as a basis for comparison. (Hsieh and Shannon 2005). At the same time, items of interest that developed during review were also coded to provide a more robust understanding of the data. Both during and upon completion of the coding, the results were analyzed for the identification of overall themes that emerged from the process.

Thematic analysis is a fitting technique for analyzing district court opinions in light of the combination of isomorphism and individuality expressed in them. Judges tend to follow a similar style in ruling on motions such as the requests for exclusion studied here. They start with a recitation of the basis of the lawsuit, a summary of the motion or motions that are to be resolved, a brief statement of relevant facts and legal standards, and an application of the legal standards to the fact. As discussed in greater detail below, this is the first of several indications of cognitive simplification by judges – the creation of what amounts to a checklist or scaffolding for the legal analysis allowing reflexive rather than reflective assessment (Kahneman 2011: 20-22). In itself this isomorphic tendency among judges is not necessarily problematic, but it is suggestive of both potential problems and ultimately some approaches to minimizing the risks of this practice.

This is not to say that opinions are entirely formulaic or monolithic. Some are incredibly detailed in description, some almost unbelievably brief. For an example of the

former, Judge David Herndon of the Southern District of Illinois constructed a detailed, multi-paged overview of expert witness law to apply to a number of rulings in the multi-district litigation relating to the marketing of Yasmin/Yaz brand birth control. *See In re Yasmin and Yaz (Drospirenone) Marketing, Sales Practices and Products Liability Litigation* (S.D. Ill. 2011), 2011 U.S. Dist. LEXIS 145552. He then tailored it depending on whether treating physicians, social scientists or other particular subtypes of expert were involved. *See id* at * 12 (adding discussion of treating physicians); *cf. In re Yasmin and Yaz (Drospirenone) Marketing, Sales Practices and Products Liability Litigation* (S.D. Ill. 2011), 2011 U.S. Dist. LEXIS 145593, *13 (adding discussion specific to social scientists); *In re Yasmin and Yaz (Drospirenone) Marketing, Sales Practices and Products Liability Litigation* (S.D. Ill. 2011), 2011 U.S. Dist. LEXIS 145989 (lacking either). By comparison, Judge Marsha Pechman of the Western District of Washington managed to resolve an expert challenge in an order totaling four paragraphs – yet still retaining an overview of the parties, the nature of the challenge, the law, and its application. *Oswalt v. Resolute Industries* (W.D. Wash 2012), 2012 U.S. Dist. LEXIS 190123.

Other distinctions exist, too, within this general paradigm. Some announce the outcome of the motion in the very title of the document – “Order Granting Motion to Exclude” for example – whereas others place it in the opening paragraph or hold the result for the very end of the document. *Dasho v. City of Federal Way* (W.D. Wash. 2015), 101 F.3d 1025, 1027; *cf. Fuentes v. Miller* (N.D. Ind. 2015), 2015 U.S. Dist. LEXIS 82978, *AgriGenetics v. Pioneer Hi-Bred International, Inc* (S.D. Ind. 2010), 2010

U.S. Dist. LEXIS 12003. Most attempt to project a tone of seriousness, but sometimes the personality of a judge will shine through in wordplay. Perhaps the clearest case of this in this study came in a case involving a dispute over agriculture chemicals of all things. Magistrate Tim Baker described the each of the parties in *Agrigenetics* as “turning its parasite–eliminating expertise toward eradicating the other side’s experts” and in denying the motions to exclude finding “[t]he parties’ experts have proven more resistant to attack than seed corn parasite.” 2010 U.S. Dist LEXIS 12003, * 2, 16. But these differences are secondary to the overall points of commonality.

Looking at these common points as a means of understanding the decision making process revealed several dominant themes. The first was that, consistent with the Advisory Committee notes to Federal Rule of Evidence 702, exclusion of an expert is the exception, while admission is the rule, across all types of expert and cases. This may not sound surprising, yet it breaks with prior research, suggesting at least the potential for some realignment of norms among judges in the intervening time. The second is that judges do not appear particularly comfortable assessing scientific and technical methods directly, and would rather rely on the judgment of third parties or the collective judgment of prior cases where possible. This reveals itself in deference to industry standards and accepted practices without much critical assessment. In this study for example fire inspection standards and medical testimony on “differential diagnosis” were routinely accepted without much apparent scrutiny. This, along with how courts tend to sort objections and process cases, suggests a third theme - that there are a significant number of cognitive shortcuts at work when judges decide these matters. None of this should be

taken to suggest judges are not paying attention, but rather that in keeping with the fourth theme the focus is on role boundary policing and reigning in extreme conduct only when necessary. Collectively, they show that judges generally seek to keep the consideration of esoteric information about methods and expert testimony bounded within the norms of the adversarial process.

Individually each of these findings is understandable. Several of them are consistent with the idea of judicial restraint in the broad sense of not unduly interfering with what is at heart an adversarial process. Many judges report reluctance to interject themselves into a case by appointing independent expert witnesses, for example (Cecil and Willging 1993: 5, 20-21). Only a minority traditionally have seen themselves as innovators rather than as followers of the law (Rowland and Carp 1996). However, while these point to a desire to honor legal norms about their role expectations and the overall process, it also raises the question of whether judges are actually willing or even able to conduct the sort of robust assessment facially required under *Daubert* and *Kumho*. This in turn suggests that perhaps some re-alignment of either the law or the expectations of judicial capacity in keeping with Geyh (2016) may be in order.

A. The Qualitative Investigation Process

Thematic analysis occupies a middle space between constrained content analysis and the minimally structured initial approach of grounded theory. (Braun and Clark 2006, Charmaz 2006). Where the former is primarily deductive in orientation, and the latter inductive, thematic analysis can be consciously directed towards either type of inquiry or occupy a space within the spectrum between them. In this instance, the approach was primarily deductive as the focus of the analysis was relatively well-

established by the research questions rather than attempting to holistically assess judicial rulings. At the same time, it was not purely deductive as the author sought to identify potential additional sources of insight in regards to the identified research questions as well as to assess the validity of the previously identified codes.

In keeping with Braun and Clark (2006), a six step process was followed. Initially, the author familiarized himself with the data by reading and reviewing the material. In this case, the author's prior familiarity with judicial rulings generally and pre-trial rulings in particular along with reading initial samples of the data stood in for a comprehensive initial review.

Second, every case was subjected to coding. This project began with a partial code book derived from prior studies but subject to augmentation. In particular, in light of the work of Buchman (2007) Dixon and Gill (2002), Merlino, Springer and Sigillo (2011) and Kim, Schlanger, Boyd and Martin (2009), several case specific factors of interest were identified. One category of codes referred to the nature of the parties as either corporate or government actors or private individuals. This follows the general view that "repeat players" tend to have institutional advantages in legal settings as expressed by Galanter (1974: 96-103). A second was the nature of the case, such as "business tort" or "patent infringement," "civil rights action" or "contract dispute", included to attempt to address the concern that Buchman's (2007: 689 n. 10) reliance on the term "tort" to identify cases might have yielded skewed results. The context of a challenge to an expert witness, such as whether it was part of a motion for summary judgment, was also coded as both Dixon and Gill (2002:281-283) and Kim, Schlanger,

Boyd and Martin (2009) suggest judges may have a strategic incentive to strike an expert when it can in turn dispose of a case.

Another set of codes was based on how judges applied *Daubert/Kumho*. Specifically, consistent with Merlino, Springer and Sigillo (2011) substantive mentions of the *Daubert* factors of testing, error rate, peer review and general acceptance, as well as indications of positive or negative commentary on these issues, were initially included. Boilerplate recitations of the entire set of factors were also noted, but as a distinct code from substantive analysis. A fifth set of codes, again drawing on the prior work, included references to the witnesses' credentials or affiliation. Finally, for each challenge whether the expert was admitted or rejected was recorded.

Following the tenets of thematic analysis, coding at this stage was intentionally over-inclusive. Any item of potential interest was coded. Examples included the identification of repeated phrases, the precise nature of every objection raised by a party to witnesses and characterizations of how the judge described the number of steps involved in the assessment of expert witnesses. Ultimately over 2000 unique codes were used, most only once or twice in the entire data set.

As the third step in thematic analysis, patterns were assessed in the process of "searching for themes." Several noteworthy patterns became rapidly apparent. Almost all judges had a formulaic approach to these rulings, with a boilerplate recitation of law that would occasionally vary followed by an assessment of particular challenges. At the same time, unlike Buchman (2007) and Merlino, Springer and Sigillo (2011) the overwhelming number of experts seemed to be admitted by all judges. Regardless of

how many steps a judge recited the analysis as having, they broadly considered whether the witness was qualified, had used an appropriate method and was offering relevant testimony.

There were also consistent themes as to how this process was carried out. Judges seemed almost compelled to discuss the credentials and relevance of expert testimony even in the absence of a challenge by any party, whereas methods were typically only assessed to the extent and in the manner challenged in the motion under consideration. In the Fourth Circuit, in particular, a recitation of a particular case - *Kopf v. Skyrn* (4th Cir. 1993) 994 F.2d 374, seemed to signal that it would be unlikely for a challenge to qualifications or relevance to succeed except in extraordinary circumstances. Rather than engage in a direct assessment of the expert's methodology, judges were generally content to rely on prior cases or third party certification (such as industry standards) where possible for indications of reliability. Judges also seemed to place tremendous confidence in the ability of the adversarial process to expose any alleged shortcomings in the testimony of the expert. Two frequent comments were that challenges "went to the weight, not the admissibility" of the expert and that any issue could be handled through cross-examination.

Two noteworthy patterns were identified that did not appear to be discussed anywhere in the literature. The first was that judges frequently upheld a challenge to an expert, but only partially. Such rulings did not result in the exclusion of the expert, but could either be a minor impediment or a significant change – and yet all prior work seemed to code the result of a ruling solely as either admitted or rejected. The second

pattern that was not evident in prior work was a concern that the expert was in some manner testifying improperly. That is, rather than an issue with their method or credentials, the expert was testifying in a manner that would usurp the function of the judge or the jury or had been improperly instructed by counsel. All of these patterns were refined into themes and influenced the development of the quantitative models discussed below.

At the same time, a negative pattern also emerged. In a break with the results of Dixon and Gill (2002) and Merlino, Springer and Sigillo (2011), the named *Daubert* factors were rarely applied meaningfully. Only 26 cases were identified as mentioning the factors as a set in this initial round of coding, and only 4 were coded as actually analyzing the full set of testing, error rate, peer review and general acceptance. There were several other mentions of individual factors, but typically only in response to a direct challenge by a party. Thus, it was determined that attempting to focus on the influence of individual factors in quantitative analysis would not be a useful endeavor. At the same time this pattern suggested an additional qualitative basis for asserting that judges are engaging in cognitive simplification processes.

The last three steps all related to the themes developed from the forgoing process. In the fourth step, the themes were reviewed to see if they were aligned both with regard to the particular codes and the data set as a whole. Fifth, the themes were identified with greater specificity. These include a theme of defaulting in favor of admissibility; a theme of relying on socially constructed understandings of expertise both in terms of particular categories of evidence and particular witnesses; a theme of seeking cognitive

simplification of the decision making process where possible, including simplifying the overall analysis, substituting more readily comprehensible considerations such as qualifications and relevance for actually assessing methods, and generally avoiding intervention where possible in favor of trusting the parties to address issues through the adversarial process; and a theme of policing the boundaries of acceptable behavior by lawyers and experts to reign in extreme behavior, rather than applying a cognitively costly analysis to most disputes. Collectively, these themes suggest that judges seek to honor legal norms about their role expectations and the overall process, but may in so doing deviate from the intent if not the letter of *Daubert* and *Kumho*. In the final step of thematic analysis, these themes were analyzed and reported with what appeared to the author to be compelling examples of each theme in the following sections in lieu of a standalone report.

Theme 1	Excluding an Expert is an Unusual Step/Admission is the Default
Theme 2	Judges Rely on Socially Constructed Standards to Assess Methods and Experts
Theme 3	Judges' Rulings Demonstrate Significant Cognitive Simplification
Theme 4	Decisions to Exclude or Limit Testimony are Largely About Addressing Extreme Cases and Policing Role Boundaries

Table 2: Themes Identified in Qualitative Analysis

B. Excluding an Expert is an Unusual Step

The literature on the impact of *Daubert* is at times conflicted. On the one hand, reading Buchman (2007: 681-85) one would get the impression that the default standard for a conservative judge is to exclude experts when challenged, whereas liberals were slightly more likely to admit an expert than not. Merlino, Springer and Sigillo (2011: 9-15) find no default to exclusion generally, but the overwhelming rejection of toxicology

experts in their sample suggests there should at least be some highly disfavored types of testimony. At the same time, neither of those reports comport with Helland and Klick's (2012: 17-31) finding that cases in both state and Federal courts have similar numbers of experts. Nor is it in keeping with Cecil and Willigang (1993: 5, 20-21), who find most judges want to respect the adversarial process rather than intervening with regards to expert witnesses. In fact, both the text of the *Daubert* and *Kumho* opinions as well as the Advisory Committee notes to the revision of Rule 702 in 2000 that codified *Daubert/Kumho* noted that exclusion is the exception, not the rule. Instead, they point to the availability of actions including cross-examination, presentation of countervailing witnesses and jury instructions as means to address weak expert witness evidence without excluding it.

If this study conformed with Buchman's findings, it would raise significant questions about the validity of the legal model. Merlino, Springer and Sigillo's findings of a differentiated standard by evidence type would potentially suggest that either the law was disregarded in some contexts or that the standard for admission for particular forms of testimony had been made more stringent by post-*Daubert* developments. The latter outcome is suggested in the work of Beyea and Berger (2001: 348-350) as well as Beecher-Monas (2011: 41-43), both of which report the adoption of a doubling of risk as a simple (and extremely difficult) threshold for toxicology evidence to cross in the first instance. In that context, while it might be bad science, such a limit on admissibility would not be probative of an extra-legal influence in the sense of the attitudinal, behavioral or strategic models of decision making. By comparison, a default of exclusion

as to all forms of testimony would be a direct rebuke of the Supreme Court's comments in *Daubert* that they were attempting to recognize a more liberal standard for scientific evidence than that allowed by *Fry*.

Ultimately, however, the present study differed significantly from Buchman (2007) in this regard. It appears that in the great majority of cases judges are at least attempting to honor the statements in the Advisory Committee notes to Rule 702 to use exclusion sparingly. Although there will be a further discussion of quantitative data and analysis in the next chapter, even a cursory look at the rate of exclusion makes this distinction clear. Where Buchman found the mean admission rate in all cases was just under 44 percent (Buchman 2007: 681), the overall admission rate in this sample was in excess of 81 percent. When controlling for region, there was some variability, with the Ninth Circuit issuing a rejection in 25 percent of cases where the Fourth and Seventh Circuit were issued a rejection in less than 16 percent of cases, but nowhere approached the frequency of rejection reported by Buchman (2007). As discussed in greater detail in Chapter 4, this likely is attributable to Buchman's search parameters, as he looked for cases that specifically invoked the word "tort" rather than looking more broadly at cases that sounded in tort with or without the word appearing *in haec verba*.

Courts do not even appear to be as hostile to toxicology evidence as suggested by Merlino, Springer and Sigillo (2011: 9-15). This may have something to do with the fact that they appear to be much rarer in recent years. The term only came up in 4 cases in the Ninth Circuit, for example, and was really only contested in one. It was similarly infrequent in the other circuits, which itself may be a residual impact of a prior degree of

hostility and learning by lawyers after the ruling on remand in *Daubert* not to rely on toxicology for specific causation. However, the mere fact an expert was speaking on issues of toxicology did not seem to foretell their rejection. If anything there has been an effort to educate judges about the nature of such testimony – the Federal Judicial Center dedicated a chapter in its Reference Manual on Scientific Evidence on the subject, for example – that leads to a greater degree of sophistication in the analysis of such testimony. *See, e.g. Taylor v. Union Pacific Railroad Co* (S.D. Ill 2010), 2010 U.S. Dist. LEXIS 96802, *12-15 (relying on Reference Manual and admitting expert); *cf. Aurand v. Norfolk Southern Railway Co* (N.D. Ind. 2011), 802 F.Supp. 2d 950, 959-962 (comparing methodology to Reference Manual standards and rejecting expert).

As a general rule, judges appear to want to admit an expert when they feel there is a plausible justification of the expert’s methodology. As will be discussed below, this often leads to a “safe harbor” effect when experts suggest they are comporting with an industry norm that can short circuit the analysis. Although judges do not often explicitly refer to this as “general acceptance” it appears to be consistent with Dixon and Gill’s (2002: 285-288, 299-300) finding that the place of the expert within their field or the law has a significant impact on admission or rejection. At the same time, even when the expert deviates from the industry practice a good explanation will likely lead to admission rather than rejection.

3600 Michigan Co., Ltd. v. Infra-Metals, Co (N.D. Ill. 2011) 2011 U.S. Dist. LEXIS 574 is a good illustration of this judicial tendency. The parties disputed whether a landlord had made a good faith effort to secure a substitute tenant and thus mitigate the

harm of termination. To try and buttress its assertions, the plaintiff hired a real estate appraiser to assess the market in the relevant time period and establish if the asking price was reasonable. The expert admitted he did not use the leading industry standard for appraising property in reaching his conclusions – which the defense seized on as the basis for attacking his testimony.

This might seem like the paradigmatic “easy case” – and admissions of variance or lack of rigor are often part of a successful challenge. However, Magistrate Judge Andrew Rodovich was willing to give the expert the benefit of the doubt when there was a plausible justification for his conduct. Here, it had to do with the disconnect between the client’s need in the case – for a general survey of market conditions – rather than a focused market based appraisal of a particular property. Given that the leading standard had no specific guidance on such a task, and that the expert used techniques such as interviews and personal knowledge typical of similar studies, the court ultimately decided that the deviation was immaterial to the reliability of his methods.

Similarly, sometimes it only takes a shift in language to make an expert move from excluded to admissible. In *Exclaim Mktg, LLC v. DIRECTV, LLC* (E.D. N.C. 2014), 2014 U.S. Dist. LEXIS 166900, Judge Louise Flanagan had previously restricted the testimony of a damages expert because she found that the theory of damages lacked an adequate causative link. That is, while the calculation might have been mathematically accurate, it relied on an assumption that all losses were attributable to “defendant's communicating certain statements to independent satellite dealers” even though there was no “evidence to show that these statements reached a substantial sector

of the independent satellite dealer market.” *Id.*, 2014 U.S. Dist. LEXIS 166900 *2. This actually happened twice, and risked the plaintiff having no damages testimony at trial.

At the time she restricted the testimony, however, Judge Flanagan permitted the plaintiff to try again. This time, it went to a “lost revenue per call” calculation. Instead of asserting it lost clients as revenue streams directly because of the defendant’s statements, this theory argued that the plaintiff had to charge less per call to existing clients as the client base eroded but its supply of marketing calls remained the same. Because this had an evidentiary link outside of the expert – the testimony of one of the owners of the plaintiff corporation – Judge Flanagan found it was admissible, subject to cross-examination.

Even where the judge finds that the core of an expert’s testimony is improper, they are often loath to exclude the witness entirely. *Dasho* illustrates just how far a restriction, rather than an outright rejection, can go in limiting a party’s case. This case involved assertions of excessive force by police officers in Federal Way, Washington. Specifically, the plaintiff was shot several times by multiple officers while wielding what Judge James Robart somewhat cryptically referenced as a “blunt tipped knife” while naked and highly agitated. 101 F. Supp. 3d at 1027.

Although he was convicted of criminal assault for his conduct with the knife, Mr. Dasho sued on the grounds specifically that the shots were unjustified and more generally that the entire encounter was excessive. He retained a forensics expert who asserted he could reconstruct the sequence of events, including the precise sequence of shots fired, in a way that contradicted the officers’ sworn statements. Mr. Dasho also retained an expert

in “police procedures” who would argue that the conduct in question was excessive. Both were challenged by the defendants, although the primary focus of their claims was the shooting narrative expert, Kay Sweeney.

Judge Robart’s opinion makes clear that either Mr. Sweeney or Mr. Dasho’s lawyers did not understand their obligations in writing an expert report. The judge had already found the initial report highly problematic:

“In its prior order, the court observed that Mr. Sweeney’s report consists mainly of a list of the physical evidence followed by Mr. Sweeney’s conclusions about shot sequence and Mr. Dasho’s movements. . . . The report contains little reasoning connecting the Opinions to the physical evidence and offers no information to support the reliability of the Opinions. . . . Based on that inadequate record, the court could not determine whether Mr. Sweeney had employed reliable methods and applied those methods reliably to the facts.”

101 F. Supp. 3d at 1031 (internal citations omitted).

Despite all these problems, Judge Robart had not excluded the witness as a discovery sanction, but requested a supplemental filing “detailing Mr. Sweeney’s methods, their reliability and how Mr. Sweeney applied those methods to this case.” *Id.* What the judge got, though, was none of that – instead, there were a number of conclusory assertions by counsel the methods were “universally accepted” without citation to any source. There was no direct explanation of the method, but instead a trial transcript that suggested the method for establishing the shot pattern was Mr. Sweeney’s subjective creation from physical evidence of a “sequence of events that seems to him to be the best overall match.” 101 F. Supp. 3d at 1032.

Not only could Mr. Dasho’s counsel not support their claims with a single, citation, Judge Robart stated that despite doing his own research he could not find “a

single instance of a federal court allowing a forensic expert to offer opinions of this type based on a like methodology.” 101 F. Supp. 3d at 1032. Instead he found several cases rejecting the idea of reconstructing a shot sequence based on such items of evidence as shot trajectories, entry and exit wounds. If that was not bad enough, it turns out Judge Robart was not writing on a clean slate with regard to Mr. Sweeney. His methods had been assessed more than a decade earlier by the Washington Court of Appeals and found to be wanting in any sort of rigor. 101 F. Supp. 3d at 1033. For these reasons he found the methods unreliable. Moreover, the judge found the methods such as they were had not been applied in a way that could be assessed for reliability – there was just no way to even assess many of the claims that were made. If Judge Robart had attempted to apply the core *Daubert* factors, the subjective assessment would obviously lack testability/replicability, any known error rate, had never been subjected to peer review and had not been accepted by any source the Judge could find much less generally within the field of forensic science. Perhaps it was for the best Dasho’s counsel did not devote much effort to even trying to justify this testimony.

The plaintiff’s police practices expert, a former police Chief named D.P. Van Blaricom, did not fare much better. He was to testify specifically as to how the use of force in the particular case compared to police standards. However, either the expert or counsel were not comfortable actually asserting that the conduct in question was more probably than not excessive force, but only that it “may have been.” 101 F. Supp. 3d at 1035. Unfortunately for the plaintiff, “may have been” standing alone is irrelevant, as it

falls to meet the standard of liability. As a result, the court excluded this element of the witness' testimony.

Even though Judge Robart had seemingly gutted the most important elements of the plaintiff's case, he didn't find that the two witnesses could not testify in their entirety. While Mr. Sweeney's shot sequence reconstruction evidence was excluded he had also provided a number of foundational statements about shot trajectories, wound patterns and locations that would be permitted to come in. Likewise Mr. Van Blaricom was not excluded in his entirety – he'd still be able to testify to general police procedures that would help the plaintiff establish the scene for their version of events and how the police departed from established norms.

This reluctance to exclude experts is understandable at several levels. It is initially consistent with a finding that judges generally adhere to the legal model in rendering their decisions. On a facial reading, it is difficult to claim anything but the law is the dominant influence in these decisions – all judges seem to generally favor some degree of admission. The Advisory Committee notes to Rule 702 are cited with some frequency, as are admonitions in *Daubert* and *Kumho* that the goal of these cases is to liberalize the admission of expert testimony. Thus, it is consistent with the precepts of the legal model that judges would bear these factors in mind and consciously choose to admit experts by default and thus follow the law.

This approach is also consistent with more general legal norms that extend beyond the four corners of the *Daubert/Kumho* line of cases. Although judges are cautioned in *Daubert* to worry about the potential for experts misleading a jury, they are

also fundamentally committed to the value of the adversarial process. That not only includes a belief in the ability of lawyers to correct the record through cross-examination and alternative witnesses, it also manifests itself in a reluctance to be seen as interfering with the parties' presentation of a case. Cecil and Willigang (1993: 20-21) reported this was a widespread reason judges would not appoint independent expert witnesses. Similarly Rowland and Carp (1996: 27-54), while finding policy making in district courts in extreme cases, also found most judges were reluctant to be seen as actively engaging in making law instead of following it. Thus it would be out of character for most judges to seize upon expert witness determinations to exclude witnesses altogether.

At the same time this trend can also be seen as at least partially driven by cognitive influences that judges may not openly admit. The statements of law in *Daubert/Kumho* that the standard for admissibility as well as the Advisory Committee notes, along with the standard of review being set at an abuse of discretion in *General Electric v. Joiner*, creates a "choice architecture" in the words of Thaler and Sunstein (2008: 6) that favors admissibility. As Baum (2006: 50-53, 99-100) notes and Kozinski ([1993]2013: 117) echoes, professional reputations matter to judges. Being reversed can be a blow to that reputation. It also typically means that the same judge will have to repeat prior efforts, which is at odds with the pressures courts are under to settle and resolve cases promptly. It is a far less risky option to admit an expert from the perspective of a judge concerned with avoiding the reversal. This may well bias judges in favor of admission in normal cases consistent with the precepts of prospect theory as set forth by Kahneman and Tversky (1979: 277-280, 289-90).

Nonetheless, as *Dasho* makes clear, being reluctant to exclude testimony in its entirety is not the same thing as being inattentive to the issues surrounding expert witnesses. It appears that judges are taking their obligations seriously, and at least consciously attempting to follow the edicts of the law to simultaneously demand intellectual rigor from experts while applying the standard for admission liberally. In doing so they are also following the expectations of cognitive theories of decision making. This does not rule out the possibility of attitudinal, behavioral, or strategic impacts also shaping judicial decision making. As discussed in the remainder of this chapter and the next, it appears that several influences matter depending on the context. At a minimum, though it is worth remarking on the fact that at least in the broadest sense there appears to be support for the idea that judges are striving to follow the law in keeping with the normative legal model. They undoubtedly fall short of that goal, but this seems less a mark of any sort of conscious political or strategic concern and more a matter of basic humanity.

C. Judges Rely on Socially Constructed Standards to Assess Methods and Particular Experts

In most models of judicial decision making, there is an implicit assumption that the judge's conduct is predicated on a thorough analysis of the case's substance. In the legal model, for example, it is taken as a given that the judge comprehends the facts and law and applies the latter to the former. Strategic models similarly assume that there is some degree of awareness of the impacts of ruling in one way or another, both in terms of the law and the potential political fallout. Although the attitudinal model asserts that judges are motivated by policy interests, and their judgment is impacted by the cues in

the case, it still asserts at a minimum the judge is fully apprised of the underlying circumstances of the case and assesses the merits directly, albeit through a lens shaped by their personal views. Behavioral models tend to break with this to a certain degree, as they assert there is something in the background of the judge that drives the decision, and this may not be a conscious influence. However, even here the assumption is that judges are acting in a reflective manner, fully versed in the facts and law of the matter before them – it is just that the assessment is impacted by their background to a degree they may not be aware of or capable of filtering out.

This assumption does not mesh well with information on how people generally assess complex questions. As Simon (1955: 101-105) pointed out in one of the earliest challenges to rational actor theory, there are very few circumstances in which an individual will truly engage in comprehensive assessment of all the information that is potentially relevant to a decision. More often they will establish parameters and adopt a solution that “satisfices” those parameters. So long as the minimum threshold is met, in most cases decision makers stay within their boundary conditions, as the cost of further information or assessment is not justified.

Prior empirical assessments of rulings on expert witnesses suggest that there is some degree of this type of satisficing being practiced by judges. Dixon and Gill (2002: 285-288, 299-300) noted that judges seemed to emphasize whether an expert’s testimony was of a type “generally accepted” in the field, despite *Daubert*’s indication that this standard from *Frye* was not controlling. Merlino, Springer and Sigello (2011: 17) similarly found that general acceptance was regularly cited. This is also consistent with

Helland and Klick's (2012:32-33) finding that admission rates appeared to be similar whether a court only applied the simple and allegedly narrower *Frye* standard or followed *Daubert*.

This tendency is understandable. Judges evaluating expert witnesses are routinely asked to evaluate a broad range of esoteric topics. In one case, they may looking at something as seemingly trivial and yet highly technical as the proper design of a sidewalk and its height relative to a street. *Delarosa v. Speedway, LLC* (S.D. Ill. 2013), 2013 U.S. Dist. LEXIS 125104. The next case in the pipeline may, like *Dasho*, require an evaluation of novel approaches to shooting reconstruction. A third dispute, such as the *In re Yasmin and Yaz* multidistrict litigation (MDL) cases, might span everything from specific and general causation to labeling requirements to medical monitoring efforts – there were 16 experts submitted on liability related issues alone in the consolidated phase of the proceedings. A fourth might raise questions about how to evaluate supposed expertise on the “culture of NASCAR” in a contract dispute over sponsorship. *See Team Gordon Inc. v. Fruit of the Loom, Inc* (W.D. N.C. 2010) 2010 U.S. Dist. LEXIS 14227. Yet another might call on the court to grapple with questions of how to handle experts on intersectionality, group pressure and implicit bias among members of traditionally disadvantaged groups. *See Apilad v. North American Gay Amateur Athlete Alliance*, (W.D. Wash. 2011), 2011 U.S. Dist. LEXIS 159575. No one could realistically be expected to understand the nuances of such a diverse array of fields of expertise in the compressed time afforded district courts to resolve these disputes.

Instead, judges frequently look to third party standards as a means of assessing the validity of the particular method being challenged. These are often drawn from industry groups or from prior legal evaluations of similar methods. As noted in *3600 Michigan Co.*, deviation from a standard is not fatal if there is a good explanation for it. Moreover, in the case of a developing technology a court might be more apt to overlook a lack of consensus or clear standard in the field. *See Ruppel v. Kucanin*, (N.D. Ind. 2011), 2011 U.S. Dist LEXIS 167505 (discussing and permitting testimony as to results of diffusion tensor imaging or DTI for diagnosing brain injury). However, where they do exist they create a “safe harbor” or “legal science” effect that can potentially distort the value of evidence and make it either unduly difficult or extremely easy to admit (Jasanoff 1995:130-131; Jasanoff 2008: 128-129; Beyea and Berger 2001: 348-360; Beecher-Monas 2011). Although there has been a movement by the FJC and the National Academies to align scientific understanding with legal requirements through a reference manual on science generally as well as specific fields of science, its relatively infrequent citation suggests that it has yet to achieve the same impact as the social construction of industry bodies or court cases.

In this study, two fields that came up with some frequency and highlight the issues of these safe harbors were fire investigation and medical causation testimony couched in terms of “differential diagnosis” or “differential etiology.” As to the former, the frequent citation of the National Fire Protection Agency’s standards reveals that judges will often defer to industry norms without much scrutiny. As to the latter, the frequent use of judicially constructed standards of “differential diagnosis” and

“differential etiology” even as the FJC manual has sought to clarify that these terms as used in law do not reflect medical standards raises questions whether courts are truly honoring *Daubert*. This problem is further underscored with the recognition in the medical community, as reflected in the latest edition of the FJC manual that statements of certainty by medical professionals should be taken with caution given the different values underlying medical care and legal testimony.

The National Fire Protection Agency is a trade group that, among other things, writes fire codes that can in turn be adopted by local governments. In 1992, it sought to create a standard for fire investigation that created a formal method (Watson 2015: 24). It was modeled on basic principles of the scientific method including observation, evidence assessment, alternative consideration, hypothesis development and testing. It was reportedly quite contentious among practitioners at the time of its adoption because of its use of formal language and the fact that it was presented as a consensus approach to fire origin investigation (Watson 2015: 25). Over the ensuing two decades regular revisions have occurred and despite complaints at times from field practitioners of investigation it has become a widely accepted industry standard.

This acceptance has in turn led to many court cases that treat the invocation of NFPA 921 as creating a default presumption of admissibility for fire investigations. Emblematic of this is *Citizens Ins. Co. of the Midwest v. LG Electronics USA, Inc* (S.D. Ind. 2012), 2012 U.S. Dist. LEXIS 1127128. There Judge Richard Young denied a motion to exclude expert testimony in a case where fire origin and causation was at issue largely predicated on finding that the fire origin expert had adhered to the tenets of NFPA

921. He began his discussion by finding that NFPA 921's general outline of investigation processes was a reliable method. He then proceeded to compare that general method to the assertions of the experts in question and found that they had generally followed the method.

What is remarkable about this discussion is just how general of a description of an investigation process was seen as sufficient to deem NFPA 921 a reliable method by Judge Young. Citing the 2008 edition of NFPA 921, he noted:

Chapter 4 outlines a basic scientific methodology known as the "systematic approach" to investigating fires (*Id.*, § 4.1). "With few exceptions, the proper methodology for a fire or explosion investigation is to first determine and establish the origin(s), then investigate the cause: circumstances, conditions, or agencies that brought the ignition source, fuel, and oxidant together." (*Id.*). Such investigation requires an examination of the scene, interviewing witnesses, and testing the results (*Id.*, § 4.3.2). The empirical data collected, which is "based on observation or experience and is capable of being verified," is subject to an analysis premised upon inductive reasoning (*Id.*, §§ 4.3.3, 4.3.5)

2012 U.S. Dist. LEXIS 112728, *7

Essentially, this is a basic overview of inductive reasoning and nothing more. It does not suggest why or how the investigator can be assured that their results are valid from following this generic process. It offers no epistemological standards to ensure that an investigator's assessment is truly reliable. Yet, because "[c]ourts throughout the country have recognized that NFPA 921 offers a comprehensive, peer-reviewed, and detailed guide for fire investigation, and have held that its methodology is reliable for purpose of Rule 702" Judge Young followed suit without further examination. *Id.*

Similarly, in comparing the outline of the method to the challenged opinions, Judge Young seemed willing to indulge a very broad application of the method by the

expert. Stephen Cottingham testified that he followed NFPA 921 by making observations, conducting interviews and constructing a hypothesis predicated on his process. However, his testing of the hypothesis was hardly what might be seen as rigorous in most fields – in particular, he “exchang[ed] ideas with other investigators on a more hypothetical level – i.e. ‘how can we explain these facts to either support of to no support a [] given theory?’” 2010 U.S. Dist. LEXIS 112728, *14 (brackets in original, internal citation omitted); *see also* *Severn Peanut Co., Inc. v. Industrial Fumigant Co.*, (E.D. N.C. 2014), 2014 U.S. Dist. LEXIS 34507, *6-7 (rejecting challenge to expert methodology based on NFPA 921’s endorsement of thought experiments as adequate process to assess hypotheses). Nonetheless, this was deemed to be conforming to an adequate degree by Judge Young, and he dismissed the failure of Cottingham to perform a rigorous “depth of char” analysis – something he acknowledged the NFPA itself labels “‘the most reliable for evaluating fire spread’” and “‘the key to generating reliable data’” – as an oversight that could be explored on cross-examination. 2010 U.S. Dist. LEXIS 112728, *13-14.

This is not to say NFPA 921 is not actually reliable. In its entirety it is over 400 pages long and includes a number of very specific guidelines such as the “depth of char” analysis. However, its invocation appears to have become talismanic for many judges. This underscores how, in furthering the admonition in *Daubert* and the Advisory Committee Notes to Rule 702 to default in favor of admission, judges are potentially avoiding an actual assessment of the methods employed by particular experts. This at least creates the potential for the creation of a “legal science” subfield that gives

unwarranted weight to techniques that are neither as certain nor as widely accepted as the testifying expert in an adversarial setting represents (Jasanoff 1995:130-131; Jasanoff 2008: 128-129; Beecher-Monas 2011: 60-62).

Indeed, at times parties have attempted to make NFPA 921 an exclusive source of reliable fire related investigation. To date, judges have resisted these attempts using different approaches. In *State Farm Fire & Casualty Ins. Co. v. Electrolux Home Products* (N.D. Ind. 2012) 2012 U.S. Dist. LEXIS 188434, for instance, Magistrate Judge Roger Cosbey rejected a challenge to the testimony of Ronald Parsons on the risk associated with a dryer fire in part by finding that although Parsons did not overtly claim to follow NFPA 921 his methods were compliant with the general process outlined in the standard. This approach was ultimately endorsed by Judge William Lee in adopting the magistrate's report. *State Farm Fire & Casualty Ins. Co. v. Electrolux Home Products* (N.D. Ind. 2013) 2013 U.S. Dist. LEXIS 84550. By comparison, in *Affiliated F.M. Ins. Co. v. LTK Consulting Services, Inc.*, (2014 W.D. Wash), 2014 U.S. Dist. LEXIS 53203 Judge Robart distinguished between cases in which an expert asserted they were following NFPA 921 and was found to have deviated from it and the idea that NFPA 921 is an exclusive standard for all fire investigation. There, the fact the expert did not claim to follow NFPA 921 and was an electrical engineer rather than a professional fire investigator made the alleged deviation irrelevant. The NFPA has thus created a safe harbor but not the only one for fire related testimony. Nonetheless this speaks to the risks of oversimplification and creation of "legal science" courts run when they endorse a particular method as inherently reliable.

A second example of this phenomenon can be seen in the area of medicine, and in particular the process of establishing specific causation through the testimony of physicians. Generally, in a case predicated on exposure to a potential causal agent the plaintiff has to establish both general causation – that is, is a causal chain possible under any set of circumstance – and specific causation – that is, do the particular circumstances related to the plaintiff’s injury support a finding of causation. Physicians are often used to establish this second step. *See Koho v. Forest Labs, Inc* (2015 W.D. Wash). 2015 U.S. Dist. LEXIS 180860. In fact, under some circumstances – such as a very close proximity in time or a lack of plausible alternative causes – a physician’s testimony on likely specific causes might satisfy both general and specific causation. *See McClellan v. I-Flow Corp.*, (D. Ore. 2010) 710 F. Supp. 2d 1092 (finding that joint damage was so close in time and so unique relative to installation of a pain pump as to render distinct evidence of general causation unnecessary).

Koho illustrates why physicians have become a favored source of expert testimony on specific causation, particularly in the context of mass torts such as product liability cases. This case was one of a number of individual cases brought against the makers of Celexa for failing to adequately warn patients and physicians about the potential risks of the drug, and in particular the risk of increased suicidal ideation. While not a class action, it was consolidated in the Eastern District of Missouri for seven years as part of MDL proceedings. In that time, one of the key common questions was general causation, which the MDL court found could be established through the expert testimony of David Healey. The same expert then submitted a tailored version of his expert report

to Judge Lesnik on remand that “briefly discusses Mr. Ilich’s case” and “concluded that the drug ‘contributed significantly to [Ilich’s] death by suicide” among other statements. 2015 U.S. Dist. LEXIS 180860, * 3 The basis for this finding was primarily accomplished by considering and eliminating another drug predicated on his view that there was not a “strong relationship with suicide” and the other drug, and the timing of the prescription of Celexa and the ultimate suicide. 2015 U.S. Dist. LEXIS 180860, *11. Judge Robert Lasnik ultimately found “Dr. Healy's methodology sufficiently reliable for his testimony to be admitted at trial” and in turn predicated a denial of summary judgment for the defendant based on this finding. 2015 U.S. Dist. LEXIS 180860, *15; *Koho v. Forest Labs, Inc* (W.D. Wash 2015) U.S. Dist. LEXIS 46050.

Although Judge Lasnik noted that an utter failure to explain a conclusion would have supported the rejection of a claim predicated on differential diagnosis, the opinion makes clear that it is largely deferring to professional judgment. Essentially, because Dr. Healey is an expert in the impacts of SSRI drugs and their links to suicide, his statements that he did not see the same linkage with the other drug as he did with Celexa and that the timeline allowed him to rule out the deceased’s depression as the most likely cause were taken at face value. It also takes the validity of differential diagnosis as a general method largely for granted, predicated on prior cases both within the Ninth Circuit and from outside in the particular context of suicides allegedly triggered by certain anti-depressants.

Judge Lesnik is far from alone in this - as with NFPA 921, differential diagnosis has become shorthand for a presumably acceptable form of testimony. At the outer limits

it is possible for a physician to be disqualified despite asserting to follow differential diagnosis, but typically only if the expert asserts they follow a standard that it can be shown they have deviated from without explanation. *See Messenger v. Norfolk Southern Railway Co* (N.D. Ind. 2015). In *Messenger*, for example, Judge William Lee rejected expert testimony asserting a plaintiff's carpal tunnel syndrome was work related and granted the defendant's motion for summary judgment. There the expert attested to have followed the American Medical Association's standards for work related injuries, but admitted in his deposition to failing to adhere to them without apparent justification. Short of that sort of "smoking gun" judges are generally willing to admit medical testimony where the expert asserts they engaged in a differential diagnosis, with any missing factors reserved for cross-examination rather than serving as a basis for rejection.

Contrast this largely deferential standard with the legally constructed standard for toxicology testimony to be relevant in many courts after the Ninth Circuit ruling on remand in the *Daubert* dispute. There, for evidence to establish a substance was more likely than not the cause of harm, toxicologists were regularly required to show a doubling of risk (Beyea and Berger 2001: 348-360; Beecher-Monas 2011: 41-43). As critics noted, this had the effect of sharply reducing the use of otherwise valid science in court. Here, by comparison, the same sort of legal construction may be honoring the intent of *Daubert* to make admission the rule, but it may be doing so at the risk of admitting less than rigorous applications of medical science.

This possibility is underscored by the fact that the legal construction of differential diagnosis and differential etiology are not generally what medical science is

focused on or how the field actually uses the terms. As the most recent edition of the *Reference Guide to Scientific Evidence* makes clear, a quest for a specific causal agent such as referenced in *Koho* is not how the term differential diagnosis is actually used in medicine:

In the medical context, by contrast, differential diagnosis refers to a set of diseases that physicians consider as possible causes for symptoms the patient is suffering or signs that the patient exhibits. By identifying the likely potential causes of the patient's disease or condition and weighing the risks and benefits of additional testing or treatment, physicians then try to determine the most appropriate approach—testing, medication, or surgery, for example.

(Wong, Gostin and Cabrera 2011, 690-91).

Similarly, actual treating physicians are generally not looking at matters through the lens required by the law. In a case like *Koho*, where the litigant is deceased, “[t]estifying . . . requires making judgments that physicians do not ordinarily make in their profession, making these judgments outside of physicians’ customary patient encounters, and adapting the opinion in a way that fits the legal standard.” (Wong, Gostin and Cabrera 2011, 694). Given that this is so, the socially constructed and widely accepted standards for physician testimony might well be creating a gap between the testimony of medical experts and the actual basis for their expertise.

Moreover, this legally constructed standard allows experts to testify with arguably greater certainty in a legal context than would be supported in their professional roles. As Kuhn (1996: 8-10) postulated, the nature of scientific knowledge, and really knowledge generally, is not a linear accumulation but rather a revolutionary process that contains periods of general consensus (“normal science”) punctuated by shifts in

understanding that are rarely accepted at first. The Reference Guide reflects this shift in medical decision making over the last several decades, grappling with prior views about the certainty of medical reasoning and shifting towards recognition of the inherently probabilistic and uncertain nature of many diagnoses. Yet the legal construct of differential diagnosis and standards such as “more likely than not” or even “reasonable medical certainty” do not track either the actual conduct of physicians or the emerging understanding in the field of inherent uncertainty (Wong, Gostin and Cabrera 2011 691-92). Rather than attempt to reconcile this split, multiple cases reviewed cited the legal definition supplied by the Reference Guide without acknowledging the fact that the Reference Guide sees a sharp distinction between the conduct of physicians in practice and the practice of physicians in court. This is true despite the overarching goal of *Daubert* and *Kumho* to ensure that expert testimony is linked to how an expert would analyze similar problems in the context of their non-legal endeavors. Thus, consistent with Jasanoff (1995: 130-131) medical testimony has taken on a life of its own and become a socially constructed standard that may fit the needs of the law but is not an accurate reflection of the state of medical practice.

Once more, this should not be taken as an assertion that any given expert is offering improper testimony or that differential diagnosis as used in court may not be a reliable method. However, what it does point to is how an understandable desire by judges to simplify the burden of assessing witnesses can lead to undermining the central premise of *Daubert* over time. Rather than simply assuming differential diagnosis is acceptable because it has been accepted in the past, under *Daubert* judges are supposed to

be availing themselves of the full scope of sources such as the Reference Guide and seeking to ensure that witnesses are truly offering medically sound testimony. However, unless appellate courts come to require such efforts or district courts come to see their roles as extending to a more robust analysis of even “generally accepted” methods, the gap between how medicine is practiced and how it is testified about in court is unlikely to be bridged.

An example of how this sort of misalignment might be avoided and yet still tend to favor admissibility can be seen in a different context in *Suchanek v. Sturm Foods, Inc* (S.D. Ill. 2015), 311 F.R.D. 239. There, in a class certification dispute, Judge Nancy Rosenstengel considered the admissibility of a consumer study conducted by the plaintiffs’ expert. Unlike the legally constructed standards of differential diagnoses and medically reasonable certainty, surveys are evaluated based on general research design principles. As the court put it “[c]onsumer survey evidence must comply with principles of professional survey research in order to be admissible.” 311 F.R.D. at 246. Judge Rosenstengel then proceeded to assess the purported defects in the survey not against a purely legal standard, the industry specific guideline or the assertions of the parties. Instead, she was able to use the principles laid out in the Reference Guide overview of survey evidence written by Professor Shari Seidman Diamond as a neutral, non-litigation motivated benchmark. The ultimate conclusion was also admissibility, but without the lingering doubt that either NFPA 921 or the reliance on purely legally constructed standards might create. However, this represented a truly unusual use of the Reference Guide in the present study.

In addition to relying on potentially inaccurate constructions of various methods the judges in this study also relied on the judgment of their peers in prior cases as to particular experts. That is, individual experts with either a positive or negative reputation in prior litigation tend to be treated in the same manner in subsequent cases. This was most notable in MDL cases, such as the *In re Yasmin and Yaz* litigation as well as suits against Novartis and Forest Labs, but it emerged in other contexts as well. Here too, this raises some question as to whether district judges are furthering the intent of the *Daubert/Kumho* trilogy with regards to examining methods rigorously and therefore raising questions about whether the legal model is descriptively accurate as applied to these determinations.

As already noticed, by the time Judge Lasnik considered Dr. Healey's testimony in *Koho*, he was not writing on a clean slate. Dr. Healey had been approved to testify as to general causation in the Eastern District of Missouri. The potential impact of this on Judge Lasnik is unclear, but it – and the fact that Forest Labs did not challenge that finding before him – likely had some impact. Similar prior treatment evidence abounds. As already discussed, in *Dasho* the shot pattern reconstruction technique had been previously ruled against in state court in a non-precedential case that the district court cited in part as a basis for justifying exclusion of that portion of the expert testimony. In *Parker v. Smithfield Packing Co., Inc.* the allowance of competing experts on time work studies was rooted in part on prior admission rulings in other cases. *Parker v. Smithfield Packing Co., Inc.* (E.D. N.C. 2010) 2010 U.S. Dist. LEXIS 102441. Similar experts appeared in cases relating to human resources practices, police practices, fire

investigation, life planning and damages. Where there was a clear tendency to admit or reject an expert, the same practice was generally followed.

Even where judges claim not to be influenced by prior rulings, it is not always clear how they could possibly avoid it. In *Aurand v. Norfolk Southern Railway Co.*, Judge Philip Simon considered the testimony of Richard Lipsey, a toxicologist. Although Judge Simon ultimately predicated his rejection on the fact that the plaintiff's report was so conclusory as to create an "analytical gap requir[ing] the exclusion of Dr. Lipsey's testimony" before he got to that issue he noted an oversight in the recitation of the witness' history as a proposed expert. *Aurand v. Norfolk Southern Railway Co* (N.D. Ind. 2011), 802 F.Supp. 2d 950, 958. Specifically, while claiming he had universally been accepted as an expert, he'd actually been rejected in the following colorful terms in Florida district court:

[T]his is the worst example that I have seen in my 15 years experience in the courts as a judge demonstrating what's wrong with expert testimony in our courts of law. It's obvious you can get a Ph.D. to say anything, and this one is prepared to say anything, gratuitous or otherwise...It's just sheerly his opinion -- and I do put that in quotes because, in my opinion, it's absolutely pure and simple, unadulterated speculation, guesswork, just blown-in. There's no scientific basis for any opinion that he has rendered, and I would consider him to be a false expert; and, therefore he will not be permitted in this court, unless the Eleventh Circuit directs otherwise.

Id. 802 F. Supp. 2d at 954 (quoting *Williams v. Orkin Exterminating Co.*, Cause No. 3:95CV30511-LC, DE 275, pp. 55-56).

It is difficult to credit that Judge Simon truly set aside such a prior view after having given it such prominence.

No expert illustrated the impact of prior testimony more clearly, though, than Suzanne Parisian. A former FDA regulator and founder of a consulting business

“specializing in matters involving the regulation of United States products by the FDA”, Dr. Parisian was discussed by judges in every court cluster considered. *In re: Yasmin and Yaz*, 2011 U.S. Dist. LEXIS 145593, *47-48. In *Bryant v. Wyeth* Judge Thomas Zilly went through her history both in the context of suits involving hormone replacement therapy and more generally. On at least one subject the only justification offered for ruling the testimony admissible was that other courts had let it in, and the primary driver for deciding to reserve judgment on another subject was the existence of a split in other court’s treatment of the subject. *Bryant v. Wyeth* (W.D. Wash 2012) 2012 U.S. Dist. LEXIS 190912, *11-19. In the Yasmin/ Yaz litigation, Judge Herndon used prior endorsements of her method in part to reject a challenge to her methods. 2011 U.S. Dist. LEXIS 145593, * 50. In Novartis related litigation in the Western District of North Carolina, Judge Graham Mullen decided to reject a challenge to the entirety of her testimony based on finding that the subject matter of the case was more similar to cases where she’d been admitted than to those in which she’d been rejected. *Lemons v. Novartis Pharmaceuticals Corp.*, (W.D. N.C. 2012) 849 F. Supp. 2d 608, 613-14 (Stating “[t]his Court finds it persuasive that all but one Court, *Hogan*, that considered Dr. Parisian's testimony in an Aredia® or Zometa® case found her testimony to be admissible as a general matter.”) Dr. Parisian thus is a body of law unto herself, distinct from the methods she employs – and with at least one split of authority on whether to admit her.

It is easy to understand why judges would want to be aware of an expert’s prior conduct. Disclosure of prior testimony is a requirement under Federal Rule of Civil

Procedure 26. At the same time, district court rulings are not binding precedent, and in theory at least a *Daubert/Kumho* assessment should be focused on the application of the expert's methodology in the present case rather than their general credibility. This focus on what other courts think of an expert raises the prospect of the creation of a favored and disfavored caste of experts based primarily on their prior treatment. As in Dr. Parisian's case, the same person can hold both a favored and disfavored status depending on the particular nature of the case under consideration. Countering that impulse is difficult but not impossible – Judge Herndon specifically noted criticisms of Dr. Parisian's manner of testimony or whether she was “uncontrollable and unhelpful” as a witness were irrelevant to his decision at the pre-trial stage. 2011 U.S. Dist. LEXIS 145593, * 52-53. If nothing else, though, the repeated reliance on the opinions of judicial peers underscores that the admission or rejection of a particular expert – especially one with a reputation – may not always rise and fall solely on the consideration their methodology as envisioned by *Daubert*.

Ultimately, it seems that judges cannot always divorce themselves from the human impulse to seek assurance in convention. This reassurance can come from a broad range of sources, from trade associations to authoritative rulings to the non-binding opinions of their peers. Each creates at least the potential for judges to minimize their own efforts in assessing proposed experts. Similar to most such efforts, there is a potential risk involved in such simplifications. In particular, it may lead to a process that is at one legally compliant and yet ultimately undermines the intent of the *Daubert/Kumho* trilogy.

D. Judges' Rulings Demonstrate Significant Cognitive Simplification

From the earliest work of Simon (1955) and Tversky and Kahneman (1974) there has been a growing recognition that judgment is frequently driven by reference to cognitive simplifications. As with Shapiro's (1965) argument about judicial incrementalism, this body of scholarship demonstrates that people, under conditions of limited information, often adopt shortcuts or rules of thumb – frequently referred to as heuristics - in lieu of more complicated thought processes. These shortcuts can be quite beneficial in many circumstances. However, at times they demonstrably deviate from classic rational choice – for example, the representativeness heuristic refers to the tendency of people to make relational judgments that ignore the base rates of various choices and overestimate unlikely outcomes. A series of studies has shown that judges are prone to several of the most well-established heuristics in experimental settings, such as representativeness, anchoring and framing effects (Guthrie, Rachlinski and Wistrich 2007; Rachlinski and Wistrich 2017: 211-220).

What has been less clear is whether the process of writing opinions might counteract the impulse to heuristic behavior. The use of checklists as a counter to oversights in fields such as medicine has been well documented, suggesting there would be a similar value to written opinions. Rachlinski and Wistrich (2017: 223) make such an argument on the theory it might force the judge to think about their decisions and reflect on whether an initial impression was accurate. That is certainly a sentiment echoed by appellate judges who speak of the impact of their colleagues' scrutiny as a restraint on any desire they might have to skirt the law (Kozinski [1993]2013: 116-117). At the same

time, some studies have suggested that legal decision making processes create rather than counter heuristic tendencies.

For example, Beebe's (2006) study of trademark law suggested how the legal field might actually facilitate a potentially problematic series of shortcuts. The nominal legal question for a court in such a case is to determine whether two competing marks create a potential for consumer confusion. Every court announces this is a conclusion drawn from the "totality of the circumstances." At the same time though, there is no national consensus on what is relevant to this determination. Instead, case law typically sets out a series of considerations that can be used in an appropriate dispute, while asserting none are dispositive. This in turn leads to courts tending to fixate on certain factors to the near exclusion of others. Some factors, when analyzed, will almost always favor the plaintiff, others the defendant. This in turn could lead to a significant loss of uniformity in legal doctrine and even individual judges adopting patterns that favor a particular type of party or cause while remaining within the law.

A somewhat similar, but possibly even more radical pattern emerged with regards to the expert witnesses in this study. As noted previously, judges demonstrated a wide degree in variability of how much space they devoted to legal standards, but the vast majority of opinions cabined the law to a distinct section setting forth a framework before addressing particular challenges. Judges typically adopted a preferred framework and then applied that boilerplate language moving forward. Most prominently in these passages – some as long as several pages, others as short as a single paragraph – would

be some recitation of the various elements of the *Daubert/Kumho* line of cases and/or Federal Rule of Evidence 702.

Typically these were supplemented by local circuit precedent. It was in this context that a phenomenon similar to that of Beebe (2006) emerged. While all courts acknowledge the controlling authority, they differ in assessing what this actually means in terms of the process a court must undertake. The number of issues to be decided, and in what order can differ. Moreover, certain circuit precedent suggests that only a very limited basis exists for rejecting proposed testimony – all of this might suggest a basis for circuit based variability in terms of outcome. It also documents one of the ways that judges significantly simplify the cognitive burden of evaluating potential expert testimony.

In its simplest form as commonly expressed in the Fourth Circuit, judges assessed the gatekeeping function as being a matter of assessing only two general factors – reliability and relevance. *See Earp v. Novartis Pharmaceuticals Corp* (E.D. N.C. 2013), 2013 U.S. Dist. LEXIS 129910, *7-8 (“Courts have distilled the requirements of Rule 702 in two crucial inquiries: whether the proposed expert's testimony is relevant and whether it is reliable.”). There is no definitive ordering of these considerations in these cases, and consideration of the expert’s qualifications is a component part of reliability. In contrast, the courts in the Ninth Circuit will use the same basic question, but add a threshold requirement of assessing the qualifications of the expert. *Bisson v. BNSF Railway Co* (E.D. Wash 2015), 2015 U.S. Dist. LEXIS 182242. In the Seventh Circuit, by comparison, this same inquiry is often (but not universally) stated as three distinct

steps, to be taken in the specific order of qualifications, reliability and relevance. *Sann v. Mastrian* (S.D. Ind. 2012), 2012 U.S. Dist. LEXIS 9107; cf. *In re Yasmin and Yaz (Drospirenone) Marketing, Sales Practices and Products Liability Litigation* (SD. Ill. 2011), 2011 U.S. Dist. LEXIS 145593, *7-8 fn. 3 (noting the different formulations, but concluding “*Chapman* simply combines the first two steps described in *Ervin* as a single test of reliability, whether the analysis is described as a three-step or two-step process does not substantively change the Court's analysis.”)

It is possible that this sort of small change in turn impacts how courts in different parts of the country treat the same question. That potential is underscored by the fact that in the Eastern District of North Carolina, most cases recite a pre-*Daubert* ruling, *Kopf v. Skyrn* (4th Cir. 1993) 994 F.2d 374, which makes some interesting statements about how to assess qualifications and relevance. This case was not universally cited throughout the dataset originating in the 4th Circuit, but it was essentially the default in the Eastern District. This case makes clear that challenges to both credentials and relevance are disfavored, which might help explain why there were significantly fewer cases in this court cluster than the other two studied.

Regardless of the impact of variations of phrasing, the general scope of the framework used in virtually all cases is to vastly simplify the task in front of the judge. The value of boiler plate as a labor-saving device is shown by how they were re-used both among judges in the same district and by the same judge in different cases. For example, Judges James Dever and Louise Flanagan in the Eastern District of North Carolina used some of the exact same phrasing in their handling of cases, beyond simply

applying the same body of precedent. *See Earp v. Novartis Pharmaceuticals Corp* (E.D. N.C. 2013), 2013 U.S. Dist. LEXIS 129910, *7-8; *cf. SMD Software, Inc. v. Emove, Inc* (E.D. N.C. 2013), 945 F. Supp. 2d 628 at 636-37. In some extreme cases, a judge's application of boilerplate extended beyond the recitation of a legal framework to the substantive analyses of issues. In both *C.A. v. AMLI at Riverbend, L.P.* and *U.S. Automatic Sprinkler, Co., v. The Reliable Automatic Sprinkler Co.*, for example, Judge Sarah Barker of the Southern District of Indiana used the exact same phrasing to dismiss an objection, clearly recycling her effort from the prior case to the latter. *See C.A. v. AMLI at Riverbend, L.P.*, (S.D. Ind. 2010) 2010 U.S. Dist. LEXIS 86530, *10-11; *U.S. Automatic Sprinkler* (S.D. Ind. 2010) 2010 U.S. Dist. LEXIS 29456, *7.

In addition to generally reducing the cognitive burden of writing each decision, the application of the general framework revealed that judges seem to favor the familiar tasks of assessing the qualifications of witnesses and relevance as tasks. Assessing witnesses generally in terms of their competence and background is something judges do regularly as part of making credibility determinations. Similarly, relevance and the proper scope of testimony are questions that judges rule on in every case and rooted in well-established legal principles all lawyers know. By comparison, assessing methodology is unique to *Daubert/Kumho* proceedings. This last category requires far more cognitive effort to engage directly and meaningfully. This is the type of situation in which the bounded rationality findings of Simon (1955: 101-105), the work of Kahneman (2011) on heuristics and biases and the insights into juror behavior of Cooper and Neuhaus (2000: 168-171) all suggest people will tend to take action where they feel

confident in their expertise and tend to find a way to avoid direct engagement with the more complex task.

Assessing the qualifications of experts is very similar to assessing the capacities of witnesses in a variety of other circumstances, including competency and credibility. Out of the three general types of assessment judges frequently cited in their legal frameworks, they actually looked at credentials in the vast majority of cases (over 190 separate entries). Even when the expert's qualifications were not challenged, the opinion frequently recited them at some length. This arguably might be a result of conditioning to preserve the record on appeal – but it also is suggestive of a judge who is otherwise unfamiliar with the subject of testimony being able to find their footing in the expert's credentials.

At the same time, merely possessing stellar academic or experiential qualifications is not sufficient for a witness to be admitted. Indeed, if anything a misfit between qualifications and proposed testimony seemed to be most likely when the expert was highly qualified - just not for what they were trying to say. For example, two separate experts - a distinguished architect in one case, a workplace safety and human factors design expert in the second - were excluded from testifying to alleged design defects in parking lots for the simple reason that the attorneys offering them did not even bother to point to any direct knowledge of traffic design on their part. *Estate of Myers v. Wal-Mart Stores, Inc* (E.D. N.C. 2011), 2011 U.S. Dist LEXIS 39164; *Roberts v. Menard, Inc* (N.D. Ind. 2011), 2011 U.S. Dist LEXIS 44628. If anything, judges are more likely to overlook a seeming deficiency in academic credentials than to be blinded

by a surplus of them – as with a fire inspector who never took a university course in engineering and did not acknowledge the industry standard for fire investigation but had been in the business of investigating dryer fires for a number of years and followed the general outlines of the standard. *State Farm Fire & Casualty Ins. Co. v. Electrolux Home Products* (N.D. Ind. 2012) 2012 U.S. Dist. LEXIS 188434.

Assessing relevance is another area that judges confront in every case. As noted above, relevance is presumed in the Fourth Circuit, for the simple reason that an expert provides some insight in most cases unless they are in no way adding to the knowledge of lay jurors. While not explicitly stated in this manner in the Ninth or Seventh Circuits, the net effect is the same. A lack of relevance is possible, but generally only if the issue the testimony was meant to address has been resolved or eliminated in some way. It can also occur when the information the expert would convey – such as compliance with a legal standard that is fundamentally different than the matter at issue in the current case – is seen as lacking any connection to the ultimate issue in dispute. It is worth noting in this context that one of the most controversial applications of *Daubert* – the limitation of toxicology causation testimony to cases where a toxicologist was willing to state a causative link at a “more likely than not” (or greater than doubling of background risk) – really was not a methodology decision but a relevance one based on California’s causation standard (Beyea and Berger 2001: 348-360; Beecher-Monas 2011: 41-43). Barring this sort of situation, judges can and do summarily find the testimony relevant.

The remaining area examined under the framework is the reliability of the expert’s testimony. Nominally, this is at the heart of *Daubert/Kumho*, with a goal of

ensuring that expert testimony reflects intellectual rigor consistent with the way an expert would behave outside of the context of litigation. Yet most of these cases have almost no discussion of methods. Less than half the cases reviewed (92) contain a recitation of the steps taken by the expert to arrive at their conclusion, much less a meaningful analysis of the methods employed.

Moreover, although *Daubert* is frequently taught in law schools for the four factors it set out for evaluating scientific testimony - testability/replicability, error rate, peer review/publication and general acceptance – they feature far less prominently in practice. Only 26 cases were coded as containing a recitation of the four core *Daubert* factors in the substantive portion of the challenge, and even fewer actually analyzed these factors as a group. The reason for this is in part tied to how *Daubert/Kumho* framed the utility of the factors and how circuit courts applied this. Somewhat like the factors in trademark doctrine, the “*Daubert* factors” were expressly couched as nonexclusive list. In expanding the gatekeeping function to all testimony, rather than what the courts deemed “scientific” testimony, *Kumho* noted the flexibility inherent in deciding which factors were relevant would be linked to the facts of a particular case.

However, in a striking difference with trademark law and other fields, courts have generally not sought to augment the list of factors. Rather, it is more typical for both appellate and district courts to find ways them to eschew the factors altogether. This is typically achieved by drawing a line between scientific and non-scientific testimony, despite the intent of *Kumho* to ensure a similar form of gatekeeping is applied in all cases of expert testimony. This divide in turn allows the judge to assess methodology against

essentially any standard – which in turn may be something that appears like no standard at all.

Two cases that are emblematic of this are *Bisson* and *Parker*. In *Bisson*, the plaintiff alleged that he had suffered a work related injury in part because of the condition of his work truck. He tendered an expert who concluded, without much discussion of methods, that there were a number of problems with the truck that placed it in a condition to cause this injury when the truck hit a pothole. Yet when the defense pointed out that these conclusions were not the product of a method that would pass muster under the *Daubert* factors, Judge Robert Whaley brushed aside the challenge:

Defendant argues that the methodology used by Mr. Curtis does not meet Fed. R. Evid. 702 nor the factors outlined in *Daubert* to help courts determine reliability. This argument, however, misses the point that not all expert testimony is based on scientific foundation. *See Kumho Tire*, 526 U.S. at 150 ("Engineering testimony rests upon scientific foundations, the reliability of which will be at issue in some cases In other cases, the relevant reliability concerns may focus upon personal knowledge or experience.")

Here, the Court finds that reliability in this instance depends on the expert's knowledge and experience more so than the methodology or theory behind any science. *See United States v. Hankey*, 203 F.3d 1160, 1169 (9th Cir. 2000) ("The *Daubert* factors (peer review, publication, potential error rate, etc.) simply are not applicable to this kind of testimony, whose reliability depends heavily on the knowledge and experience of the expert, rather than the methodology or theory behind it.").

2015 U.S. Dist. LEXIS 182242, *7-8.

In other words, sometimes so long as the expert is testifying to things within their expertise, there would be no way to exclude the testimony. If that were taken to its

extremes, it would be hard to see how any non-scientific testimony, as that term is defined by courts, would ever be subjected to much scrutiny.

Parker reflects a similar conclusion albeit phrased differently. In this case a putative class of employees asserted Fair Labor Standards Act violations linked to their employer's refusal to pay for time spent on various work-related tasks. Their expert's primary method was observational and involved simple measurements of particular tasks to arrive at an average time for each task. The defendants asserted that the expert's non-statistical method was non-replicable and violated prevailing industry standards for standard setting. Magistrate Judge James Gates rejected this assertion in colorful terms:

In advancing many of their arguments, defendants advocate a rigid application of the factors bearing on the reliability of a methodology listed by the Fourth Circuit in *Tunnell v. Ford Motor Co.*, 245 Fed. Appx. 283, 287 (4th Cir. 2007), namely, error rate, controls on application, peer review, and general acceptance. But the Supreme Court and the Fourth Circuit itself have instructed that the determination on reliability should be a flexible one. "In making its initial determination of whether proffered testimony is sufficiently reliable, the court has broad latitude to consider whatever factors bearing on validity that the court finds to be useful; the particular factors will depend upon the unique circumstances of the expert testimony involved." *Westberry*, 178 F.3d at 261 (citing *Kumho Tire*, 526 U.S. at 150-52). The circumstances here, including the nature of the subject matter as the donning and doffing of equipment as opposed to some intangible and esoteric biological or chemical process, do not demand the rigorous level of scientific precision for admissibility that defendants advocate.

2010 U.S. Dist. LEXIS 102441, *18-19.

Again, the impact of this read broadly is to give lower courts carte blanche to use or refrain from applying the *Daubert* factors in any setting not fitting that particular judge's concept of scientific testing. That in turn allows judges to tremendously reduce the cognitive burdens of evaluating a *Daubert/Kumho* issue. In particular, they can take

action that they can reconcile with legal doctrine without having to engage with the complexities of understanding methodological standards, much less applying them in a broad range of different fields.

Another way that judges seem to simplify their task is tied to their fundamental belief in the adversarial process. In *Daubert* itself, judges were admonished to remember that any expert testimony would still fall within the adversarial system. The thinking goes that cross-examination can be used to expose weak foundations, that counter-testimony may be applied and that jury instructions can be used to correct misimpressions. Notably, this has been shown to be inaccurate experimentally – juries use many criteria beyond the actual testimony to assess the witness, and not infrequently overlook critical gaps in unfamiliar or technical testimony (Vidmer 1995: 122-123, 150, 171-173; McAuliff, Kovera and Nunez 2009: 252-253; McAuliff and Duckworth 2010: 495-497; Levitan 2017: xiii-xxii). Nonetheless, it is a central premise of the training of judges and lawyers, and judges generally report a strong belief in the ability of the adversarial process to mitigate potential harm (Cecil and Willging 1993: 18-21).

The cases reviewed in this study suggest that this perspective underlies many of the rulings judges arrive at, and tends to make them skeptical of the need for intervention. On 135 occasions judges remarked on the ability of cross examination to be used to address alleged deficiencies in the methods or conclusions. This was one of the two most frequently cited bases for rejecting a challenge. The only one that exceeded it, and was frequently cited along with it, was the concept that a particular challenge went to the weight to be given the evidence, rather than the admissibility of the expert. These

concepts serve as simple shortcuts for judges to avoid wading into contested questions about methods between seemingly qualified experts. At the same time, though, while it is nominally compliant with *Daubert/Kumho* as a matter of legal doctrine, it arguably undermines the central goal of this line of cases.

Similarly, it is worth noting that one common form of cognitive simplification that escaped evaluation in this study was simple avoidance. Because the focus of this study was assessing different models for decision making, cases which did not reach a decision on the *Daubert* issues were excluded. Judges frequently noted the presence of a *Daubert/Kumho* motion but then postponed its consideration for a variety of reasons. Sometimes it was because other issues had rendered the challenge moot, because it was not yet ripe, or because the matter would be a bench trial. In any event, this tactic certainly allowed them to put off consideration of complex questions in the hopes of a subsequent resolution.

As a whole, these simplifications are consistent with findings in other areas of law and teachings from psychology and behavioral economics. Not only are district judges confronted with decisions involving uncertainty, they are also under extreme time constraints. It is only logical that they would seek to minimize the cognitive burden of their tasks. Moreover, these simplifications do not seem to lead to rulings that are non-compliant with legal norms and doctrine. Nonetheless, their existence and persistence suggest that they may lead to sub-optimal results both because they permit arguably improper testimony to go forward in cases and/or permit “battles of the experts” that ultimately confuse jurors. Eliminating these tendencies would likely require either

altering legal doctrine to limit the possibility or fundamentally shifting legal norms to encourage judges to engage rather than avoid taking a position on issue related to expert testimony. Neither is likely to be forthcoming in the near term.

E. Decisions to Exclude or Limit Testimony are Largely About Addressing Extreme Cases and Policing Role Boundaries

Between the default presumption of admissibility, the reliance on socially constructed standards and the cognitive simplifications shown above, it would be easy to conclude that *Daubert/Kumho* has no meaningful impact as a matter of law. Indeed, that was the general conclusion of both Buchman (2007: 687-689) and Hellend and Klick (2012: 32-33) in their exclusively quantitative analysis of outcomes. However, close reading suggests a less dismissive conclusion, especially once the significant numbers of restrictions judges impose are considered. What emerges is a system in which judges default to admissibility for a number of reasons, and reserve their harshest sanctions for extreme cases. They also seem to be very concerned with ensuring that expert testimony is not used to blur role boundaries, either by having experts speak beyond their base of knowledge or operate in a manner that crosses what they see as the appropriate limits of such testimony. This is consistent with survey findings originally reported by Gatowski, et al (2001: 443-445) that judges believed *Daubert/Kumho* impacted their approach to expert witnesses – it just is not necessarily as focused on methodology as *Daubert* itself envisioned. To use the metaphor of *Daubert/Kumho*, there is a gate, but it is generally open so long as the expert (and often the lawyers proposing the expert) have adhered to the broad contours of the schema associated with expert testimony in the American legal system and appear to be within the bounds of a given field.

Initially, although judges default in favor of admission, that does not make admission inevitable in all contested cases. It is hard to argue that admission is not a default – over 80 percent of the orders subjected to thematic analysis permitted experts to testify to some degree. At the same time, placing restrictions on testimony was quite common. In fact, when the restrictions are added to eliminations, judicial action of some kind occurred in the majority of cases. Thus, it does not seem to be a matter of judges being reluctant to act at all where they see a meritorious argument, but a definite preference for action short of the sanction of complete elimination.

As discussed in Part B above, this tendency to admit experts is consistent with both the legal model and the cognitive influences created by some of the controlling law. The tendency to constrain or modify expert testimony might initially suggest a deviation from these concepts. However, the *Daubert/Kumho* line of cases and the Advisory Committee comments to Rule 702 simultaneously stress liberalizing the standard for admission and ensuring judges keep a watchful eye on the potential for misleading or otherwise improper testimony. Moreover, those same norms of the adversarial process that restrain judges from taking what they see as too radical of an intervention in a case also stress that judicial officers have a critical role in ensuring the process is fundamentally fair. There is no similar choice architecture pushing judges to allow particular subparts of testimony the way there is urging them to avoid total exclusions. There is also far less risk of reversal from excluding a small portion of testimony based on a claim it speaks to a legal conclusion and needs to be stated differently than there is

in rejecting a witness altogether. Thus, this differentiated behavior still follows the law and the precepts of cognitive behavior.

At the same time, this raises the question of what could drive judges to exclude a witness altogether. Here too, both the legal model and cognitive theory suggest similar frameworks. The legal model suggests that judges will follow *Daubert/Kumho's* exhortation to be mindful of the risk of misleading or improper testimony. This would be a somewhat subjective assessment but there are a number of guideposts in the *Daubert/Kumho* line of cases as well as circuit court interpretations suggestive of when to reject testimony. Prospect theory suggests that in extreme cases individuals will court risk rather than entertain what they see as a certain loss. As applied to courts, this might indicate courts are more likely to court the risk of reversal when confronted with an expert so problematic that they see admission as riskier than rejections. In practice, judges reject what they see as extremely flawed experts on a number of independent bases.

The *Daubert/Kumho* cases urge courts to ensure that the testimony is reliable and relevant. As discussed in Part D above, courts have generally broken this into three broad considerations – is the witness qualified, based on their education, experience and similar criteria; is their method reliable in the legal sense both as a concept (typically but not always based on indications of general acceptance in a field) and as applied in the case (primarily a function of comparing claimed methods to the tasks undertaken by the expert); and is the testimony relevant when compared with the relevant legal standard. Judges typically take each assertion individually rather than holistically. They also tend

to look at each expert individually although they occasionally make relative judgments between similar experts.

A number of examples were found in the dataset for rejections predicated on each of these considerations. Qualifications were routinely challenged, but interestingly this was most successful when there was a mismatch of credentials and assertions rather than direct attacks on someone for having failed to specialize in a particular field or taken particular college classes. Assertions were routinely made that both methods generally were unreliable and that the expert did not apply them accurately – and extreme cases would sometimes lead to rejection, particularly where the expert admitted to either lacking rigor entirely or deviating from their asserted methodology in practice. Relevance challenges similarly succeeded most often where the proposed expert either framed their report or admitted that their testimony was so ambiguous that it could not definitively support a claim for relief or defense.

As discussed previously, it is unusual for experts to be rejected for inadequate but relevant credentials. In no small part that is because experience can often fill any gaps left absent by formal education. In *State Farm Fire & Casualty Ins. Co. v. Electrolux Home Products*, for example, the fact the inspector had not formally studied NFPA 921 or fire ignition was more than made up for by a combination of years of experience, professional training and familiarity with the particular dryers being analyzed. *Dasho* likewise permitted police procedure testimony driven primarily by the experience of a former police officer. Similarly the regular citations to *Kopf* in North Carolina cases

make clear that a broad range of bases exist to find someone qualified, and that only in the absence of any valid basis should an expert be found unqualified.

Yet nonetheless rejections linked to qualifications do happen. Most commonly, it seems to occur when the retained expert's background and experience is somehow disconnected basis from the conclusions they are making. As noted previously, both *Estate of Myers* and *Roberts* illustrate this disconnect when it comes to something as seemingly pedestrian as the design of a parking lot. Both the architect in *Estate of Myers* and the human factors engineer in *Roberts* were eminently qualified in their respective fields. There are likely persons in those fields who have knowledge of the design parameters and considerations relating to parking lots. However, these particular experts – or more specifically the counsel who retained them – failed to provide any explanation for how these particular individuals and their qualifications had any insights beyond those of a lay person on how to build a commercial parking lot safely.

In both these cases and several others in the same vein, the gaps were essentially confessed by the proposed witness in some way. In a patent case involving flame retardant fabrics, for example, a chemistry professor was found to lack the qualifications to assess how a “person of ordinary skill in the art” would read the relevant terms. *Precision Fabrics Group, Inc. v. Tietex Int'l Ltd* (M.D. N.C. 2015) 2015 U.S. Dist. LEXIS 4853. The expert, one Charles Wilkie, actually admitted that he did not know what the relevant skill in the art would be and adopted it from discussing the technology at issue with his colleagues. Compounding that problem, as Judge Thomas Schroeder put it, “Dr. Wilkie opined on the level of skill that a person of ordinary skill in the pertinent

art would have -- and then candidly acknowledged that he lacks this skill.” *Id.*, 2015 U.S. Dist. LEXIS 4853, *10. In particular he set the standard as requiring work with not just the chemistry flame retardants – an area within his expertise – but specifically the application of those retardants to textiles. This doomed him as a relevant expert, as his own experience was nonexistent in this field.

Rejections based on qualifications often are compounded with other deficiencies or problems. Emblematic of that is a case involving how much odorant (the chemical marker used to alert people to the presence of natural gas) a plant was releasing. *Fernlund v. Tanscanada USA Services Inc. d/b/a Transcanada* (D. Ore. 2014) 2014 U.S. Dist. LEXIS 159386. The complaining party put forth David Robinson as a witness on the question of plant operations. The first problem was that although he was an engineer who had overseen the building of fueling systems, he had no actual experience with the odorant process or the equipment used for it. If that was not bad enough, his report was characterized as vague and misleading about both what he was relying on and what the basis for his claims were. Several statements were also deemed irrelevant. Moreover, it turned out the expert was also the spouse of the plaintiffs’ counsel - a fact not disclosed in the report and only brought to the court’s attention by opposing counsel. All of this led Magistrate Mark Clarke to reject his testimony, actions that were ultimately upheld by the district court.

Successful challenges to the methods employed by experts also tend to fit a pattern of extreme conduct by the proposed expert. As discussed in Part C above, judges generally do not reject proposed methods that have been embraced by an industry or

blessed by the court. Sometimes they arise from conventions in the field not directly related to litigation such as survey methodology or the “yardstick” method for assessing values through comparison to comparable businesses or properties. Others are part of a recursive dialog between courts, lawyers and experts that can result in unique, judicially constructed safe harbors for expertise that break with typical practices, such as the judicially modified and endorsed forms of “differential diagnosis” and “differential etiology” currently used in many court cases. Still others, like NFPA 921, are the result of conscious efforts by industries to standardize practices in a manner that has clear implications for litigation but can also serve other purposes. Regardless of their origin, typically where there is a consensus in the field judges seem reluctant to second guess it. To place it in Kuhnian terms, if a field appears to the judge to be in a state of “normal science,” judges tend to accept the field’s representation regarding what counts as a valid method.

However, sometimes an entire method proves to be groundless. This is easier when the practitioners of the method are unique, as in *Dasho’s* bullet pattern reconstruction. Similarly an entrepreneur trying to carve out a unique practice “as a ‘litigation consultant’ in personal injury and wrongful death cases, wherein his only role is to shorten or lengthen the plaintiff or decedent’s life expectancy, depending on whether he has been retained by defense or plaintiffs’ counsel, respectively” had his unique method torn apart in *Fisher v. United States* (W.D. Wash. 2010), 2010 U.S. Dist LEXIS 108237, *4. Although the proposed expert, Scott Kush, held both a JD and MD from Stanford, he had no interest in practicing either medicine or law. Instead, he researched

life expectancy and had published several studies in that field. On paper he appeared to be qualified – and he was retained by the Federal government to state an adjusted life expectancy for the victim of a fatal accident in which the government was liable.

Whatever the merits of that research might have been, what he did in court cases fell far short of the expectations of the scientific method. Dr. Kush would calculate the impact of various medical and demographic factors on baseline life expectancies in a manner he admitted was not testable or replicable. He had never subjected his technique to peer review and he could not calculate an error rate. Although he claimed it was generally accepted, several other witnesses specifically contested that claim and he provided no evidence to back up the assertion. In short, while relying on scientific credentials to give his testimony credence he was essentially telling courts to trust that his subjective view was inherently correct in calculating seemingly precise reductions in life expectancy from the generally accepted U.S. Life Tables. Unsurprisingly, Judge Benjamin Settle declined the offer and rejected Dr. Kush's entire line of testimony.

Although Dr. Kush was somewhat unique in this dataset in terms of his clear efforts to create a novel, apparently litigation centered form of expertise, his treatment by Judge Settle was of a piece with how judges treated what they see as truly exceptional cases. As with the expert in *Fernlund*, *Fisher* and other cases where an expert's central arguments were rejected, judges regularly noted multiple independent grounds for eliminating an expert. There was frequently a problem with not just the method but the relevance of the evidence or the process by which the expert had arrived at their conclusions. In fact, judges in these situations often seemed to go out of their way to

document the shortcomings with the expert in significant detail. It often is not enough to find the expert is problematic – they must be documented as utterly lacking in value. Moreover, in doing so judges frequently apply the literal factors *Daubert* in circumstances that would not seem to demand them, rejecting the cognitive shortcuts used in the vast majority of cases.

This tendency is exemplified in the curious example of an expert testifying about cleaning the barrels of muzzle loading black powder firearms. *Hartman v. Ebsco Industries*, (N.D. Ind 2013) 2013 U.S. Dist. LEXIS 2013 140474. The plaintiff alleged that a manufacturing defect in a conversion kit that was designed to make it easier to use synthetic pellets as an alternative to black powder was to blame for an unexpected discharge of the weapon that had caused him injury. His expert, one Steven Howard, testified that the fundamental issue with the kit was that it lacked an ideally shaped cleaning jag – “a tool that screws onto the end of the ramrod, and is used to sweep the inside of the barrel after firing a muzzleloader.” *Hartman*, 2013 U.S. Dist. Lexis 140474, *23. “[B]ecause the jag supplied by the Defendants does not perfectly match the recessed portion of the 209 breech plug, latent embers are likely to remain after discharge of the rifle and cause accidents, such as the one involving the Plaintiff.” *Hartman*, 2013 U.S. Dist. Lexis 140474, *23-24. The expert had designed and built an alternate jag he believed was superior, but had not actually tested it in practice.

There were multiple levels of problems with this testimony. The most obvious one was that the plaintiff had not bothered to swab the barrel between multiple shots on the day of the accident. This alone should have created a basis for rejecting the expert’s

testimony as irrelevant. Yet Judge Springmann did not end the discussion there. In contrast to numerous cases that suggested the *Daubert* factors were irrelevant to practical experiments and experience based testimony, here the judge specifically explored the lack of testing, peer review and a known error rate as independent bases on which to reject this testimony. While these conclusions are not wrong in isolation – and indeed the ruling was affirmed on appeal – it seems to be a break with the overall thrust that the *Daubert/Kumho* standards are to be flexible and not applied unduly rigidly. How peer review would have been conducted or even been relevant to establishing the value of this testimony, for example, is something the judge never quite addresses.

At the same time, this behavior is precisely what we might expect from a judge who is aware they are going against the default presumption and taking the potentially riskier path in terms of possible reversal. This sort of extraordinary case might well lead a decision maker to build as complete of a case as possible to document their action. Similarly, in *Ruppel*, the evolving view of DTI in the relevant community led the judge to engage more fully with the *Daubert* factors, albeit in a setting more closely approximating *Daubert's* own fact pattern. In both situations, the judge appears to have been moved by the parties and the situation to go against the default conduct of treating the named *Daubert* factors as irrelevant to a given case.

Another of these defaults that may be contravened in extraordinary cases is the tendency of judges to defer to industry standards. As discussed in Part C it is not uncommon for an industry to create a certification or standard like NFPA 921. When followed those are generally deferred to without much examination. However, there is a

latent danger in this tendency that is demonstrated by the counter example of *Cascade Yarns, Inc. v. Knitting Fever, Inc.*, (W.D. Wash. 2012) 2012 U.S. Dist. 15097. In this long running fight over the appropriate labeling of blended fiber wools, Cascade Yarns and others in the industry had investigated their rival's claims to make a cheaper yet high quality cashmere wool blend. Consistent with industry practice, they sent a purported cashmere blend to something called the "Cashmere and Camel Hair Manufacturers Institute" (CCMI), who in turn referred the sample to an eminent bioengineer named Kenneth Langley for "fiber content analysis." *Cascade Yarns, Inc.*, 2012 U.S. Dist. 15097, *4. Professor Langley used the current versions of quantitative and qualitative testing standards developed by the American Association of Textile Chemists and Colorists ("AATCC"). The qualitative test, which purported to permit species level discrimination of fiber content, "includes detailed observation of individual fibers under a microscope both in transverse section and elongated, including observation of the scales on animal fibers." *Cascade Yarns, Inc.*, 2012 U.S. Dist. 15097, *14. He reported similar results when other industry participants sent him the same wool for sampling, suggesting the test was reliable at least in terms of replicability. In short, if the normal defaults of deferring to industry standards and general acceptance were followed admitting this testimony would be an easy call for Judge Ricardo Martinez.

In this case, though, there was a critical problem with relying on Dr. Langley. As it turned out, the CCMI sent reference samples for testing to its membership using the AATCC methods every two years. Those tests showed that before the samples that triggered the litigation were sent to Professor Langley he had performed abysmally in the

tests. He could generally (but not always) identify a pure sample of cashmere or yak wool, but frequently misidentified the nature of complex strands. Sometimes he claimed a more complex mix than was actually present, other times he saw single species in hybrid mixes. Critically, he had found no cashmere in multiple reference samples with documented cashmere levels similar to the purported makeup of the Knitting Fever samples. In other words, it was entirely possible that the linchpin of the plaintiff's case – the lack of cashmere fibers – was a baseless accusation.

Moreover, Professor Langley was not alone in his failings – the method itself was unreliable. Forty laboratories participated in the 2007 round of the CCMI's trials in 2007, all following the AATCC protocols. The results varied wildly in their estimates of blended fibers, both in terms of basic composition and specific makeup of the samples. In fact, "[o]f the 200 tests conducted (5 samples x 40 laboratories), 29 tests were accurate within three percentage points of the actual value for the fibers (wool, cashmere, yak, and angora)." *Cascade Yarns, Inc.*, 2012 U.S. Dist. 15097, *19. In fact, the errors were so widespread that Professor Langley's estimates approximated the sample average in several cases, something the plaintiff's counsel actually tried to use to save his testimony, along with relying on the expert's long history and general acceptance by his peers. Although Judge Ricardo Martinez refrained from openly mocking the plaintiff's counsel, he did archly refer to the former effort to as an attempt at "statistical sleight of hand" and had no problem in rejecting the expert as well as bringing to light the problems of the broader industry.³

³ Interestingly, although not directly on point with the subject of this study, several years later the FBI admitted their own method of microscopic hair analysis was fatally flawed,

While Judge Martinez in this one instance was able to overcome the bias in favor of industry standards and other generally accepted measures, the extraordinary evidence it required raises troubling questions. There is no comparable objective reference for all the practitioners of the general outlines of NFPA 921. As the FJC Reference Manual on Scientific Evidence makes clear, the standards of differential diagnosis and differential etiology in court do not track the processes doctors use in treatment. Moreover, there is a growing recognition that diagnosis, as a complex and probabilistic exercise, can have a high degree of error and subjectivity – doctors often make presumptive diagnoses that are ultimately inaccurate, and frequently multiple doctors disagree over how to interpret the same data (Wong, Gostin and Cabrera 2011:691-694; Baker 2005: 502-503). This doesn't mean experts should be rejected as a matter of course, as that would suggest only testimony free from uncertainty should ever be admitted. At the same time, it does suggest more care should be paid to the process by which industry standards or conventions were arrived at and whether there are any meaningful benchmarks to ensure the methods actually are reliable in the sense of *Daubert/Kumho*.

The *Cascade Yarns* case is also emblematic of why judges likely have such faith in the power of the adversarial system. The CCMI's test results were obtained through discovery. They were used effectively by Knitting Fever's counsel to a degree Judge Martinez was moved to take the extraordinary step of eliminating an otherwise eminently qualified witness and indeed rendering a whole category of evidence suspect. From the perspective of a judge (and likely Knitting Fever), the system worked as designed – the

with erroneous statements in more than 90 percent of cases reviewed by the FBI and the Innocence Project (FBI 2015).

testimony was examined and found wanting on the merits through the tools of adversarial proceedings.

Similar, albeit sometimes less dramatic examples were common in the dataset. When experts were rejected entirely it was often attributable to their own testimony or conflicts between that testimony and various assertions of counsel as to what the expert had done or would say. Thus, John Lamberth, a social psychologist who had discussed the difficulty of accurately assessing the racial disparities in law enforcement from passive observation limited to traffic violations, nonetheless adopted these techniques in a racial discrimination case and claimed they were valid – a claim Judge Thomas Schroeder found lacking in credibility. *United States v. Johnson*, 122 F. Supp. 3d 272 (M.D. N.C. 2015). It did not help that the methods of single observer, fixed observations Professor Lambeth had relied on to reach his opinions were actually criticized in the very articles he pointed the court to as supporting his techniques, and that he essentially relied on subjective understandings of visual indications of Hispanic ethnicity with no apparent effort at meaningfully testing those approaches. All of this combined to lead to the exclusion of his testimony, which in turn made it virtually impossible for the Federal government to maintain its assertions of racial discrimination.

A similar fatal admission was made by Stephan Neese, a purported accident reconstruction expert in the matter of *Stachon v. Woodward*, (N.D. Ind. 2015), 2015 U.S. Dist. LEXIS 129958. The key question in the case was the location of a pedestrian when they were struck by a truck. Neese argued that the impact was between 100 and 130 feet from the final resting place of the plaintiff based on the location of a single sock. Yet he

admitted under cross examination that clothing was an inappropriate marker and had no explanation for why he thought one sock was reliable but another located more than 60 feet away was not. This, when coupled with the fact his “reconstruction” used this assumed distance, an unrepresentative truck and a wooden model that scaled out to a 6’9” human and an impact distance 60 times greater than observed led Magistrate Judge Andrew Rodovich to reject the testimony in its entirety.

Extreme examples like this and similar situations likely cause judges to overestimate the value of cross-examination and contrary testimony. Judges are trained from law school and practice to think of cross-examination as the key skill at impeaching the testimony of witnesses, limiting or destroying their value. Similarly, counter witnesses are stressed as giving the jury a choice and allowing them to accept the side they find most credible. When judges are moved to rule it is often because of these tools and how they highlighted particular gaps in testimony, both in the context of expert witness and elsewhere.

This in turn may trigger the cognitive shortcut known as the availability heuristic. As Tversky and Kahneman (1974: 1127-1128) pointed out, generally speaking frequent events are easier to recall, and thus what is most readily recalled can often be relied on as an initial estimate of likelihood. However, when unusually salient or exceptional events are judged against those that are more mundane, the exception may be thought of as more likely to occur than the mundane event because of the ease of recall. For judges, their entire careers involve looking for and ruling on questions contested through the adversarial process. Instances of highly effective cross-examination or arguments where

the judge had confidence in the correct nature of their ruling are part of every judge's career and are apt to be highly salient in light of a judge's training and experience. Thus, it is easy for a judge to think that these same tools will be highly effective if and when a jury is exposed to them. This may explain the prevalence with which judges note that an objection goes to the weight to be given evidence by the jury rather than its admissibility and the ability for the matter to be explored through cross-examination.

While these sentiments are thus understandable, empirical studies suggest cross-examination and contrary witnesses on technical matters are not very efficacious in helping jurors arrive at an accurate understanding of issues (McAuliff, Kovera and Nunez 2009: 252-253; McAuliff and Duckworth 2010: 495-497). Initially, witnesses often cannot spot significant errors that would raise validity concerns about testimony, with and without cross-examination. When subject matter is so complex as to be beyond juror's knowledge to process, their reaction is often to focus on peripheral issues such as how much the witness is paid or their connection to the district as a means of assessing their general trustworthiness (Cooper and Neuhaus 2000: 168-171). This limits the possibility that cross-examination on technical questions is apt to be effective at either educating jurors or swaying them. Moreover, the introduction of a rebuttal witness who noted methodological questions was not to lead jurors to evaluate the study more carefully, but to be skeptical of all expert testimony in a "pox on both your houses" effect (Levett and Kovera 2008: 369-370). The latter impact might benefit defendants, as the neutralization of a plaintiff's expert would typically undermine their ability to support

critical elements it must establish to prevail. However, it certainly does not further the aims of the adversarial system.

Nonetheless, judges are unlikely to suddenly be disabused of the value of the adversarial system. The reviewed cases suggest that one of the key tasks undertaken by judges in *Daubert/Kumho* rulings are policing the roles of the adversarial process. Every key player in a trial has a legally defined role that in theory is supposed to be kept distinct. The judge informs the jury as to the law, in the form of instructions (supplemented but not controlled by the arguments of the parties) at the start, end and during the course of the proceedings. The witnesses supply the jury with testimony as to disputed factual questions. The lawyers craft that testimony into arguments to justify a verdict in their favor. The jury determines both what versions of facts are most credible and what conclusions are to be drawn from applying the law to the facts.

Frequently, to try and make their argument more compelling, expert witnesses' conclusions will track the ultimate question in a case – such as that a party was negligent, or that they failed in their duty to the alleged victim. It is not technically improper to embrace an ultimate conclusion under Federal Rule of Evidence 704, but it still is not seen as helpful for an expert to simply offer a conclusory and unspecific assertion about the law. Thus, one of the most common grounds for exclusion (although still dwarfed by the reasons for admission) of testimony, was the 13 times proposed testimony was found to invade the role of the court in whole or part. Similarly, attempts to speak to the credibility of a witness or otherwise invade the role of the jury led to at least partial exclusion 12 times.

Oftentimes both of these complaints are found in the same case, as an expert seeks to direct the jury on how to follow the law and what experts they should credit in making this determination. *Kennedy v. United States Postal Service* (N.D. Ind. 2013) 2013 U.S. Dist. LEXIS 41904. In *Kennedy*, a former USPS employee asserted her termination violated the Federal Medical Leave Act (“FMLA”) and regulations related to the same. The plaintiff retained a business administration professor named William Anthony to testify to his conclusion, which tracked the ultimate issues in the case – that from a human resources perspective the conduct violated the law and regulations. The plaintiff obviously tried to assert that there was a distinction between this and classically improper credibility testimony, but Judge Philip Simon found that it was a distinction without a difference. In the same case, Professor Anthony also was retained to opine that certain testimony by the defendant’s representatives was improper or evasive in some manner. Again the judge rejected assertions this somehow did not amount to testimony on credibility, stating:

Even if I accept this flat denial as gospel, which is a stretch given the nature of the testimony, his conclusions regarding Bush's motivations for answering questions in a particular way are still speculative. They are not based on reliable principles or methods. Indeed, the report fails to detail any scientific or experience-based method by which Anthony can arrive at the conclusion as to why Bush answered certain deposition questions in the manner in which she did. These opinions are nothing more than Anthony's subjective belief.

Kennedy, 2013 U.S. Dist. LEXIS 41904, *16-17.

Another line that judges are focused on policing is that between counsel and experts. One of the central premises of *Daubert/Kumho* is that an expert is not supposed to be a pure hired gun, malleable to the will of counsel. At the same time the nature of

the adversarial system almost dictates that their testimony will fit the arguments of the party that retained them. In fact, the concept of adversarial presentations is so engrained that neutral court appointed experts actually have less impact on jurors than those that are presented as being retained by a party (Brekke, et al.1991: 468-470). Thus, judges will not infrequently restrain the use of experts or testimony that they suspect is too much the mental product of counsel and too little the independent work of the expert.

As with most objections, judges are generally reluctant to strike experts on the grounds that they collaborated with experts. However, where they feel it is clear that the expert has ceded control to counsel they will reject the resulting testimony. In *McClellan*, Judge Ann Aiken's finding that general causation testimony was unnecessary was all the more remarkable because she also excluded key studies based on litigation bias. Specifically, she found a 396 patient study nominally conducted by Frederick Matsen to support the link between joint infusion pumps and chondrolysis. Although the expert claimed there wasn't a pre-determined outcome to the study, Judge Aiken found that there was simply too much involvement of counsel from research design and case selection to identifying the nurse who did the actual data extractions and providing guidance to both her and Dr. Matsen over the course of the study. With all of this involvement, Judge Aiken found this evidence was too infected with litigation bias to be presented in court under *Daubert/ Kumho*.

Cases that trigger the exclusion of witnesses are clearly the exception. In the vast majority of disputes, experts are presented who are within the mainstream of their fields, use only the most conventional of approaches and are admitted to testify in a way

consistent with those standards. At the same time, these exceptions document how judges are focused on meaningfully applying *Daubert/Kumho* to restrain excesses by parties and advocates but may be constrained by their own perception of the proper role of judges. The result of this is a focus on keeping expert's testimony within broad expectations of the adversarial process. This is a meaningful form of oversight, and in keeping with the overall thrust of *Daubert/Kumho* – but it can create a perception that the gatekeeping function is little more than *Frye* by another name. It also creates the potential for unexamined norms and cognitive shortcuts to undermine the objective of trying to ensure that expert testimony is grounded in sound methods and practices.

These findings suggest that at a minimum, judges are acting consistently with at least two different models simultaneously. In the first instance, the almost universal default in favor of admission, regardless of location, judge or subject matter, is consistent with a broad conception of the legal model. *Daubert/Kumho* as well as Rule 702 instructs judges to default to admission and honor the precepts of adversarial legalism and they obey. Judges in the Ninth Circuit deviate from their colleagues consistent with heightened consequences of potential errors in decision-making. The law has a predictive impact on judicial behavior.

At the same time, the means of accomplishing this compliance suggests that judicial conduct is also consistent with a cognitive model. As with most human endeavors, judges seek to reduce their cognitive burdens. They craft and abide by boilerplate language that makes their tasks more routine. They simplify the seeming complexity of understanding expert testimony by reducing the question to a handful of

issues. They substitute the judgment of their peers and industry for the more burdensome and risky task of supplying independent analysis in many cases. In each of these instances we can see the significant influence of cognitive processes, driven both by general human impulses and the unique conditioning judges receive about the norms of adversarial legalism. Thus, there is no single simple answer to why district judges act as they do.

Moreover, understanding the full scope of these themes cannot be achieved by qualitative assessment alone. Although these statements are grounded in clear examples, it may be that extraordinary cases create a form of availability bias and have been overrepresented. To counter this possibility, and to test the influence of factors that cannot be ascertained from the actual cases, a quantitative investigation was undertaken. That study, which is the subject of the next chapter, is generally consistent with the qualitative assessment, but identifies some factors that appear to influence these decisions that most judges would not admit openly, and suggests that judicial pronouncements about which types of challenges are favored or disfavored may not always align with actual outcomes of contested cases.

Ch. 4 - Quantitative Evidence Suggests Judges' Primary Influences are Legal and Cognitive, but Political Identity and Location Matters at the Margins

Although the qualitative analysis revealed significant evidence in support of the influence of legal norms and standards as well as the impact of cognitive shortcuts on decision making, the picture was necessarily incomplete. Judges do not include a disclosure of their biographies or overtly surface their biases in their opinions with much frequency. Establishing the potential influence of politics or behavioral conduct cannot arise effectively from the four corners of the document. Similarly, questions of strategic decision making and the potential impact of factors such as location are more amenable to examination at a reserve from the individual decision level. Therefore, a quantitative analysis was conducted to fill in this picture, as well as to test the initial impressions arising from the qualitative assessment.

As a preliminary necessary step, biographical information on the judges involved was compiled. This revealed a portrait that is more complex than is often portrayed in the literature. The population of judges is less diverse than the broader American society. Far fewer racial minorities or women are appointed to the Federal bench than their proportion of the population. Appointments tend to be uneven by location, rather than following consistent trends. Senators in several states also follow practices that stress consensus, rather than political ideology, in appointment. These factors suggest that decision making might be more isomorphic than otherwise expected. They also speak to issues that need further investigation but are beyond the reach of the present study.⁴

⁴ For example, there is apparently a more constrained path to appointment for women and minority judges than for white men, but further investigation of the judiciary would be necessary to draw any conclusions about this pattern.

The quantitative analysis of the decisions through the multiple lenses of different models similarly revealed a complex portrait of decision making. At its most basic level, this analysis confirmed the qualitative analysis in many respects. Judges appear to be primarily driven by the law, and default to the admission of experts. They demonstrate a more active form of gatekeeping once restrictions are included in the analysis. Objections that align with the conduct of judges in other fields, such as relevance and policing role boundaries, tend to be more likely to be granted. In all these ways, this appears to support both the legal and cognitive models of decision making.

Even at the descriptive level, though, the quantitative analysis suggests several additional considerations may matter to some degree. Although all judges tend to admit experts, controlling for partisanship reveals a degree of difference, with judges appointed by liberal presidents more willing to admit experts than those appointed by conservative presidents. The location of the ruling matters as well, with a higher – and somewhat counter-intuitive – rate of exclusion and restriction in the Ninth Circuit. At the same time, some expected associations with location based on qualitative analysis and prior studies were not evident. Behavioral and strategic factors also appeared to have some impact on outcomes, although the robustness of these factors largely proved to be illusory once political partisanship was introduced as a control variable. This suggested to some degree the judicial attitude model is also explanatory of judicial conduct at this level.

When all the disparate factors were considered in a formal model attempting to capture elements of all dominant models of decision making using regression analysis, the picture came into somewhat sharper focus. Nonetheless, it is not a simple portrait.

When the dependent variable is a judge's choice to restrict testimony, circuit of origin of the case and the types of objection raised have a statistically significant relationship to whether a judge acts, and the choice to restrict correlates at a statistically significant level with the choice to publish the opinion in the *Federal Supplement*. The same findings hold true when the dependent variable is whether the judge has taken some action, including both restrictions and exclusions, but the partisan alignment of the judge is also statistically significant.

In the case of both restrictions and general judicial actions, the results can be interpreted as being consistent with the cognitive shortcuts identified in the qualitative analysis as well as the legal model. Challenges that go to the methods used by an expert – subjects judges are less likely to feel confident assessing and/or would defer to cognitive shortcuts such as industry standards - are associated with lower rates of judicial action. By comparison, challenges based in the more familiar subjects of relevance and the proper scope of testimony are associated with significantly higher rates of judicial action. Judges appear to be honoring the overall thrust of *Daubert/Kumho* to liberalize the admission of expert testimony, but also maybe skirting around the core of the gatekeeping obligation as a way of avoiding difficult cognitive efforts.

If, in keeping with Buchman (2007), the dependent variable is restricted to the less frequent choice of exclusion, the picture changes somewhat. In this setting, none of the legal factors are significant. Moreover, challenges claiming a problem with the propriety of testimony flip their sign and become associated with admission rather than rejection. Partisanship and location retain their statistical significance, but the model as a

whole is not much improvement over random chance. In particular, while the finding on partisanship is consistent with Buchman, the limited improvement the addition of partisanship makes to the explanatory power of the model suggests that its role may have been overemphasized in prior studies. Moreover, the apparent significance of location varies depending on which circuit is used as the reference for multiple dummy variables.

Ultimately, these results suggest that law is the dominant influence on judges, but not the only one. The universal tendency in favor of admission and relatively small variation between circuits suggests judges are in fact primarily following the precepts of the legal model. There is room for local variation consistent with the law, although not always in expected ways. At the same time, politics matters at least as a marginal concern, with Republicans as a whole more likely to engage the extreme sanction of exclusion. It also likely becomes more important on issues of greater political salience. Additionally, how the judges apply the law is generally consistent with cognitive shortcuts, both in terms of the general admission rate and the issues to which they tend to be most responsive. At the same time, in the exceptional circumstances triggering exclusion, it appears that the impact of cognitive shortcuts is attenuated. All of this suggests that judicial decision making at the district court level is more nuanced than frequently discussed, and that further attention should be paid to both the legal model and deviations from it.

A. The Hypotheses Under Quantitative Investigation and Variables Assessed

Based on the literature as well as the results of the qualitative analysis a set of twelve hypotheses relating to the research questions set out in Chapter 2 were developed.

Each is related to one of the major models of judicial decision making set out in that chapter. Each is grounded in the prior work in the field of judicial decision making as well as the themes developed in the qualitative investigation. They are:

H1. Consistent with the precepts of the legal model, judges as a class will show a common tendency to admit all experts.

The first hypothesis is based both on the qualitative analysis as well as the statements in both *Daubert/Kumho* and the Advisory Committee notes to revised Rule 702. If in fact judges are applying the law, there should be a fairly high rate of admission regardless of what other factors such as party type, location, ideology or anything else that is considered. If a consistent rate is found similar to Kahan et al. (2016: 391-397) or those reported by Galanter (1993: 70-71) in many fields of decision making this would appear to be strong evidence for the influence of law. If, on the other hand an admission rate of under 50 percent is found similar to Buchman (2007: 681) this would suggest other influences such as ideology are potentially more significant.

H2. Judges within the Ninth Circuit will uphold challenges to expert testimony at a higher rate than their colleagues in other circuits.

The second hypothesis is primarily derived from *Mukhtar*, 299 F.3d at 1064-1066. That case holds that a failure to provide an adequate basis for a ruling on the admissibility of expert testimony can result in the invalidation of a trial rather than a limited procedure on remand to re-evaluate the expert admission question. This in turn does not directly suggest judges in the Ninth Circuit would be more hostile to expert testimony, but they may be more reflective and cognizant of the implications of their choices. That sort of “slow thinking” is potentially likely to lead to greater critique and mindfulness in keeping

with rational actor models of decision making (Kahneman 2011: 12, 21-22). That in turn suggests both that a higher than average rejection rate would be both consistent with the legal model and the broad precepts of strategic models such as the revised model of Epstein and Knight (2013) and particularly the principal-agent model of Songer, Segal and Cameron (1994: 674-675).

H3. Judges within the Fourth Circuit will be less likely than their colleagues to uphold objections based on the qualifications of an expert or the relevance of their testimony.

This hypothesis is based on another example of unique circuit-based standards. In the Fourth Circuit during the qualitative analysis there were repeated references to *Kopf v. Skyrn* (4th Cir. 1993) 994 F.2d 374. Like *Mukhtar* this case suggests a circuit difference that could influence judicial behavior but its expected impact would run in the opposite direction. In particular, *Kopf* establishes that it should be difficult to reject an expert based on their qualifications and presumes expert testimony is relevant. Logically, this should mean those types of challenges would be disfavored relative to circuit's that do not have such a clear signal in their case law.

H4. Judges appointed by Republican Presidents will be more likely to uphold challenges to expert testimony than those appointed by Democratic Presidents.

H5. Judges appointed by Republican Presidents will be more likely to uphold challenges to expert testimony made by corporate or public defendants sued by individuals than those appointed by Democratic Presidents.

The fourth and fifth hypotheses rest both on the findings of Buchman (2007: 681-685) and general views of how conservatives and liberals view expert testimony and civil litigation generally. The precepts of the judicial attitude model suggest that when judges have the opportunity to act in accordance with their policy goals there should be a

significant variation in outcomes between conservatives and liberals (Segal and Spaeth 2001). The relatively deferential treatment of expert witness admission determinations under *Joiner* provides district courts with that freedom of action, although the political salience of expert rulings is less than clear. Tort reform and hostility to the current state of civil litigation is associated with conservative values (Kagan 2001: 147-150).

Moreover, *Daubert/Kumho* was interpreted by conservatives as creating a tool that would permit the elimination of “junk science” that had been asserted as being the reason for the proliferation of a number of types of litigation. (Schwartz and Silverman 2006: 226-231). Moreover, the vast majority of *Daubert/Kumho* challenges are brought by defendants (387/510 in the current dataset) to try and eliminate plaintiff’s cases, which in theory should be in the interests of conservative judges. This is the assertion of Buchman (2007: 687) based on his study of cases including the word “tort” in them, with conservative judges being far more critical of expert testimony than liberals. This suggests that judges appointed by Republicans should be more likely to exclude expert witnesses and specifically those tendered by individual plaintiffs suing businesses or public entities.

H6. Female judges will be more likely to uphold challenges to expert testimony.

Boyd (2013: 211-213) finds that female judges have an overall pattern of more rapid settlement and case resolutions. She attributes this to the socialization of women and specifically that they are less attached to the model of adversarial resolution of disputes. Given this, and the finding of both Beecher-Monas (2011: 39) and Dixon and Gill (2002: 294-296) that the exclusion of experts is often associated with the rapid

termination of disputes, this same “problem solving” mentality could lead female judges to be more likely to reject expert witnesses.

H7. Judges on senior status will be less partisan than active judges.

This hypothesis is derived from the revised strategic model of Epstein and Knight (2013: 12-13) and specifically the goal seeking incentive of promotion. District court judges on active service are in a sense “auditioning” for elevation to appellate court, as they are one of the key sources of such appointments (Slotnick, Schiavoni and Goldman 2017: 397-399). Demonstrating a consistent ideological pattern, such as being tough on “junk science” or demonstrating empathy with individuals could be beneficial to that effort. (Posner 2008:114). Once on senior status, though, any such incentive disappears and if anything judges are obligated to moderate their opinions. (Posner 2008: 159). If in fact this holds, then senior status judges should be more similar to each other than to active judges with similar ideological profiles.

H8. Judges will be more likely to exclude expert testimony at the summary judgment stage.

This hypothesis is rooted in both the strategic model and the suggestion of Kim, Schlanger, Boyd and Martin (2009) that district court decision making is likely context dependent. One of the central goals of the modern judge is rapid resolution of cases short of trial (Boyd 2013: 197-198; Robbennolt, MacCoun, and Darley 2010: 29). A motion for summary judgment presents a district court with the opportunity to eliminate a case without the costly time of trial. Doing so based on a *Daubert/Kumho* ruling will render the case less vulnerable to reversal on appeal under the deferential standard afforded under *Joiner*. Moreover in the event of reversal the judge is no worse off in terms of

their docket load than they were before and has had the case off of their calendar for months or years. Given the suggestive finding in Dixon and Gill (2002:294-296) that motions to exclude seem to be more successful as a component of a summary judgment judges may be employing *Daubert/Kumho* rulings in this context a strategic means of limiting or eliminating the case entirely.

By comparison excluding an expert on the eve of trial is less likely to terminate the proceedings and a reversal after trial can be exceptionally burdensome. In that setting, as Rowland, Trafficanti and Vernon (2010: 183) noted pre-trial rulings remain a significant strategic opportunities for a district judge to influence the outcome of a case. However in light of the enhanced risk a judge might be more inclined to shape testimony through a restriction or instruction than exclude the expert entirely.

H9. Judges will be more likely to restrict rather than reject the testimony of an expert witness even when a legal challenge is upheld.

The American system of adversarial legalism is party driven rather than judge driven (Kagan 2001: 105-106). Judges are reluctant to intrude on the adversarial process even to appoint expert witnesses in accordance with Federal Rule of Evidence 706. (Cecil and Willging 1993: 18-21). Given this situation, and the observations of the qualitative analysis in this study, it seems likely that judges would prefer to shape and monitor testimony rather than exclude it in its entirety.

H10. Consistent with the concept of heuristics, judges will be more likely to uphold objections that are similar to rulings they are regularly called on to make.

Throughout the qualitative analysis, judges appeared to be attempting to simplify their decision-making process. Although they did not overtly rely on the general

acceptance *Daubert* factor in the manner reported by Dixon and Gill (2002: 285-288, 299-300) or Merlino, Springer and Sigillo's (2011: 17), they did often seek to avoid analyzing methodology directly in lieu of comparing the expert to prevailing industry or legal norms. The impression of the evidence was that objections rooted in technical matters were more likely to be seen as going to the weight of the evidence rather than admissibility. By comparison, questions of relevance and the appropriate bounds of testimony are common to every case and areas in which judges are likely to have developed significant expertise. In such settings both Simon (1955) and Kahneman (2011) suggest that a decision maker will focus on their areas of comfort and minimize the discomfort of addressing unfamiliar subject matter.

H11. Judges appointed before the completion of the *Daubert/Kumho* trilogy will be less likely to uphold challenges to expert testimony.

This hypothesis is again predicated on socialization. Dahl (1957: 285-291) and Keck (2016: 252-258) both note the Supreme Court and the bench as a whole is an evolving body. Bailey and Maltzman (2011:94-100), and Songer, Sheehan and Haire (2003:125-126) both demonstrate that judicial practices change over time based in part on what the new generation of judges see as the norms of behavior and shortcomings with prior conduct. Thus judges appointed under Roosevelt deferred more to Congress and less to precedent, whereas more recent appointees have demonstrated the inverse pattern of behavior. Judges who were appointed after the completion of the *Daubert/Kumho* trilogy likely took gatekeeping as an established part of their practice rather than an innovation, and have had the benefit of familiarity with the concept. To judges in service while these changes were implemented the shift was likely more jarring. As reported in

the survey conducted by Gatowski, et al (2001: 444-448) these judges did not understand the more technical elements of *Daubert*. As a result, it would be expected they would be more reluctant to adopt the new framework.

In addition to informing these hypotheses, two findings from the qualitative analysis also called for an effort at re-coding the material for quantitative analysis. The first was that there was significant variability in how judges treated serial motions to exclude. Some assessed each motion in isolation, where others would dispose of a large number of pending motions into an omnibus ruling. At the same time, in both instances judges nearly always made individualized rulings on particular experts. The rare occasions where a judge rendered a single ruling that applied to more than one expert in its entirety happened when the experts had written a common report and were intending to testify as to a single subject, such as when two accountants had created a damage model or two statisticians had collaborated on a single analysis of dryer fires. Based on this, it was determined that the unit of analysis for quantitative assessment should be the individual expert rather than the motion as a whole. This resulted in coding 510 separate opinions, 429 of which were rendered by Article III judges and 81 from magistrates.

The second revelation from the qualitative analysis that suggested a different approach to coding was the relative lack of assessment of the *Daubert* factors in most cases. Instead, it was possible to generally place challenges to admissibility and their analysis by judges into four broad categories – those going to qualifications, those attacking methods, those claiming the testimony lacked relevance and those that asserted

the testimony was improper. Individual challenges were re-coded to place them in these categories.

Each individual ruling was coded for several independent, control, and dependent variables. The independent and control variables fell into four categories – those that relate to the nature of the case in which the expert is proposing to testify, those that relate to the court and judge deciding the case, those that relate to the procedural status of the case and motion to limit the witness’ testimony, and those that relate to the specific challenges raised and if they were successful. The dependent variable under most hypotheses tested is the outcome of the ruling, which can take three distinct forms but is essentially dichotomous (although, as discussed below in greater detail, that dichotomy can be expressed differently depending on whether the question being asked is about whether the judge excluded an expert or whether a judge granted the motion in whole or part). To test the third hypothesis, though, the dependent variables are the success rates of challenges to the credentials of an expert and the relevance of their testimony, with the independent variable of interest being the circuit of origin of the dispute. Similarly, the tenth hypothesis looks at the relative success rates of different types of challenges, rather than the outcome of the case as a whole.

The first category, variables that relate to the nature of the case, includes six variables. These are the type of plaintiff, the type of defendant, which party is challenging the testimony⁵, an interaction variable built from combining plaintiff type,

⁵ For the status of challenging party, the coding is based on how the parties are initially listed as plaintiff and defendant rather than their status relative to the particular motion. Thus, even if an expert retained by the plaintiff is actually testifying in a defense of a counterclaim, and is therefore opposed by the proponent of the counterclaim, the motion

defendant type and who brought the challenge, whether the challenge was brought in the context of a motion for summary judgment and whether that motion was successful.⁶

The first four are both relevant to a partial replication of Buchman (2007) regarding the influence of judicial attitudes and an exploration of whether Galanter's (1974) observations about the relationship of "one shotters" and "repeat players" holds in this context.

Initially, the interaction terms was created to test Buchman's (2007) conclusions about the influence of party status and partisanship in the condition most likely to reveal either. Per Buchman (2007: 687), challenges brought by defendants – particularly large corporate or government defendants – should be especially favored when they are brought against the testimony supporting the case of individual plaintiffs. Finding such an empirical imbalance between corporate and individual actors would be consistent with the general argument of Galanter (1974) regarding how repeat players are in a favored position relative to individuals in litigation. Moreover, adding controls for partisanship should in theory reveal that conservative judges are favoring motions by corporate and government defendants as this would align with the objectives of the tort reform movement to limit the ability of individuals to pursue legal redress. (Kagan 2001:147-150).

is still coded as being brought by the defendant. This was done to preserve alignment between this variable and the party identification variables.

⁶ Initially, it had also been planned to use the type of expert as a variable, consistent with Merlino, Springer and Sigillo's (2011) findings of differentiated rates of admissibility between toxicologists (over 80 percent excluded) and psychologists (over 80 percent admitted). However, ultimately there simply were too few experts opining on either toxicology (10) or psychology/neuropsychology (20) to meaningfully evaluate the possibility of discipline related differences in treatment of experts.

By comparison, the co-presentation of motions for summary judgment and motions challenging expert witnesses provides an opportunity to study strategic action by judges consistent with the eighth hypothesis. If rejecting an expert is the basis for in turn finding no evidence for a necessary element of a claim or defense, it theoretically could insulate the ruling from attack on appeal. Summary judgment findings are generally reviewed *de novo*, but *Daubert* determinations are given the much more deferential abuse of discretion review regardless of when they are made. (E.g. *C.W. v. Textron, Inc.*, (7th Cir. 2015) 807 F.3d 827 (upholding grant of summary judgment based exclusively on excluding expert's conclusions as unfounded)). Dixon and Gill (2002: 294-296) reported that rates of rejection in these circumstances were remarkably higher, but also noted that it was too rare of an occurrence to evaluate meaningfully. Here, 152 of the 510 challenges were resolved in the same ruling as a motion for summary judgment. Although not every challenge to an expert overtly tied success or failure of the summary judgment motion to the fate of the challenge, this should still provide a robust enough sample to assess if judges are engaging in strategic grants of motions to exclude experts.

Party type, who is bringing the challenge and summary judgment status and result are all nominal values coded for dummy. The interaction variable is effectively a dummy as it can only be a one if an individual plaintiff is being challenged and a zero in all other situations. Table 3 summarizes the details of these variables and how they relate to the hypotheses being analyzed in this study.

Variable Name	Variable Type	Possible Values	Relevant Hypotheses
Plaintiff Type	Nominal/Dummy	0= Corporate or Public Party 1= Individual Party	H1, H5
Defendant Type	Nominal/Dummy	0= Individual Party 1= Corporate or Public Party	H1, H5
Defendant's Challenge	Nominal/Dummy	0=No 1=Yes	H1, H5
Corporate Challenge of Individual Plaintiff	Interaction of Plaintiff Type*Defendant Type*Defendant's Challenge	1=Individual Plaintiff's Witness is Challenged by Corporate or Public Defendant 0=All Other Situations	H1, H5
Summary Judgment	Nominal/Dummy	0=No 1=Yes	H8
Summary Judgment Granted	Nominal/Dummy	0=No 1=Yes	H8

Table 3: Case Nature Variables.

The second category of variables relates to the court and judge making the decision. These include the circuit in which the ruling arises, whether the decision was published or not and if it was what publication it is in, the “judicial common space” score for the judge, behavioral characteristics including gender, racial minority status, whether they were appointed pre-*Kumho*, are active or have taken senior status, and what their prior job was before being appointed to the Federal Bench.

All data related to the judges’ behavioral characteristics for an individual case was taken from a code book created to serve as a reference. In the code book, each Article III judge’s name was associated with the date of their appointment and appointing President, indications of their current status (and if they had taken senior status, the date of the change if it was during the time period studied), gender, race and prior position. For

magistrates, dates of appointments, gender and prior position were coded where such information could be verified. In total, records were compiled on 75 district and magistrate judges, as described below in Table 4.

Circuit/District	Article III Judges	Magistrate Judges	Total
Fourth	13	5	18
East NC	6	3	9
Mid NC	3	1	4
West NC	4	1	5
Seventh	20	9	29
South Ill.	6	1	7
North Ind.	8	4	12
South Ind.	6	4	10
Ninth	23	5	28
Oregon	6	4	10
East. Wa.	10	1	11
West. Wa.	7	0	7
TOTAL	56	19	75

Table 4: Judges Represented in the Sample

All the information in the code books were created from public sources.

Appointment dates, gender, race and prior work experience for Article III judges were identified through the Federal Judicial Center’s (“FJC”) Biographical Directory of Article III judges.⁷ Magistrate judges’ information was primarily compiled from the websites maintained by each of the component courts, although in two instances for magistrates within the Fourth Circuit this was supplemented with outside information from other websites. In two cases, relevant background information on magistrate judges could not be found and therefore minority status and prior occupation data was not coded and those cases were excluded where appropriate.

⁷ This directory is maintained at <https://www.fjc.gov/history/judges>. Although individual judicial listings do not contain direct indications of gender or minority status, the advanced search tools on the site permit identification of judges based on those criteria.

In compiling this information, it became apparent this group was more isomorphic than the broad parameters of American society and in some ways confounds how political views and behavioral traits like gender and ethnicity normally align. This is in part a legacy of the historic dominance of the elite ranks of law schools and the legal profession by white men. Even considering this historic legacy, the fact that only 12 of the 56 politically appointed judges in the sample were women, and only 7 were identifiable as racial or ethnic minorities (including 2 African American women, 2 African American men and 3 Hispanic men) is striking. Of all the judges in the Fourth Circuit who made rulings in this data set, only 1 of 13 political appointees was a woman, and none were minorities. Moreover, 3 of the 5 minority male judges were appointed by conservative Presidents, suggesting that if judges do track the political preferences of their appointing officer this might confound any effort at identifying a connection between views broadly shared by minorities and judicial decision making. Thus, although the behavioral model was in part assessed based on gender in this study and minority status was used as a control, any conclusion about the role of gender or minority status must be taken with caution and is at most suggestive and a basis for further exploration.

Most traits could be coded directly. However, the appointment date was used to help create a “judicial common space” score to assess the fit of the judicial attitude model with these decisions (Giles, Hettinger and Peppers 2001: 636-638; Buchman 2007: 680-81). These scores are constructed based on the first dimension of DW-NOMINATE scores relative to the time of appointment and range from a theoretical

score of -1 (for extreme liberalism on economic issues) to 1 (for extreme conservatism on the same issues). The appointing President's first dimension score is the starting point, and what Buchman (2007: 680-681) relied on exclusively. A modification derived from Giles, Hettinger and Peppers (2001: 636-638) is to account for the role of intra-party Senatorial courtesy by substituting the average of the first dimension DW-Nominate scores of the President and any home-state senators of the same party (Giles, Hettinger and Peppers 2001: 636-638). Thus, if there was a senator from the President's party in the state the judge was appointed to, that score was averaged with the President's. If both senators were from the same party as the President, the first-dimension scores for the pair were first averaged and then averaged again with the President's score. The resulting number – ranging from a -.45 for certain Clinton nominees to .73 for some George W. Bush nominees – was entered into the code book.

The interval score was used in all regression analysis. However, for certain initial analyses, the interval data was too granular to be of much use – individual variations of a one-hundredth of a point were treated as distinct cells in cross-tabulation, for example, rendering the interval of little help as a control. At the same time, as the premise of the hypotheses relating to the impact of partisanship is that Democratic appointees as a category should vary from Republican appointees re-coding partisan affiliation as a categorical variable was appropriate. (Pollack III 2011:58-59, 212) For purposes of identifying Democrats and Republicans for cross-tabulation, negative scores were recorded as 1 in a dummy variable and positive scores were 0.

It should be noted that this process of placing judges in a uniaxial space is not universally accepted. Bailey and Maltzman (2011: 41-42) have shown the value of additional data points such as “bridging observations” about how individual members of the Supreme Court compare with the attitudes of various political figures on contested issues. However, finding such information for district court judges, particularly given the low political saliency of most of the issues they confront, is elusive, leaving researchers to fall back on the less sophisticated proxies.

The current study points up the potential shortcomings of this approach. On the one hand, there are cases in which partisan alignment would be foolhardy to ignore. Michael Simon, a judge in the District of Oregon appointed by President Obama in 2011, is married to Suzanne Bonamici, a fairly liberal member of Congress representing Oregon’s First district since 2012 (Fassenden 2014). It is hard to imagine that Judge Simon does not align with liberal values in terms of his own views to some degree. Thus, for him a DW-Nominate score of $-.37$, a fairly moderate score achieved by averaging the score of President Obama and Oregon’s two Democratic Senators Jeff Merkley and Ron Wyden is sensible – and may even understate his views.

At the same time the other Obama appointee for the District of Oregon reflected in the sample defies simple political categorization. Marco Hernandez was initially nominated by President George W. Bush in 2008 (Oregon Editorial Board 2011). Had he been confirmed based on that nomination, the average of President Bush’s first dimension DW-Nominate score and Republican Senator Gordon Smith would have given Judge Hernandez a judicial common space score of $.73$, suggesting severe

conservativism. Instead, though, Judge Hernandez' score is identical to Judge Simon at -.37, over a full point lower.

It would be absurd to think Judge Hernandez went from staunch conservative to moderate liberal in a matter of months, but that is the conclusion the judicial common space system leads to. The reality is at once much less radical and yet suggestive of why any assumption that judicial conduct will align with the political outlook of an appointing President should be made with caution. The reason for this shift lies not in any change by Judge Hernandez but in the practices connected with judicial nominations in Oregon. Historically, nominations have come from a short list recommended by bipartisan committee with insight into the local community and consent of both senators (Wyden and Merkley 2017). It was on Senator Wyden's recommendation that Judge Hernandez was added to the list of nominees generated by the commission in 2009 for submission to President Obama (Oregon Editorial Board 2011). Ultimately President Obama accepted the recommendation, and nominated Judge Hernandez twice before he obtained unanimous confirmation in 2011.

Judge Hernandez' nomination is an extreme version of a not uncommon story. Similarly, when Indiana had a split Senate delegation between Democrat Evan Bayh and Republican Richard Lugar, there was a policy of consultation and consensus on appointments regardless of which party held the White House (Odendahl 2012). These norms seem to have eroded under the Trump administration, which appears to be basing many appointments on ideology without much input from local communities or Senators, but with regards to studies of current judges they interject a note of caution in assuming

there is direct correlation of DW-Nominate scores with the outlook of a given judge (Brown, 2017; Wyden and Merkley 2017).

Similarly, using first dimension DW-Nominate scores fails to fully account for the potential impact of time. As documented in Table 5 the judges who issued opinions in this dataset covered six separate administrations, with an irregular pattern of appointments. Some districts were significantly impacted by a given President with little or no impact on adjacent districts:

Pres.	WD Wa.	ED Wa.	D. Or.	9 th	SD Ill.	ND Ind.	SD Ind.	7 th	ED NC	MD NC	WD NC	4 th
Carter		1		1					1			1
Reagan	3			3		4	2	6	3			3
Bush 1					1			1			1	1
Clinton	2	2	3	7	3		1	4		1		1
Bush 2	5	1	1	7		3	1	4	2	3	1	6
Obama		3	2	5	2	1	2	5			1	1
Total	10	7	6	23	6	8	6	20	6	4	3	13

Table 5: Judicial Appointments by Circuit and Administration

As Dahl (1957: 285-291), Keck (2016: 252-258) and Bailey and Maltzman (2011:94-100) all demonstrate for the Supreme Court, and Songer, Sheehan and Haire (2003:125-126) show for courts of appeal, norms and practices shift on an intergenerational basis as the composition of the bench changes. A judge appointed by Presidents Reagan or Carter, like 14 of the judges studied here, might have a profoundly different outlook on legal norms than the 11 appointed by President Obama. While DW-Nominate attempts to capture some of this fluctuation, the first dimension score is anchored in views about economic policy that are relatively persistent over time. Thus variations in legal expectations may not be reflected accurately in these values. Nonetheless, as judicial common space scores are an accepted approach within the field

and allow for comparison with prior studies such as Buchman (2007) this system was retained for assessing the potential influence of partisanship on judicial rulings.

Again, this cautionary note should not be taken as a rejection of the concept of the judicial attitude model in its entirety. There is simply too much evidence at the appellate and especially the Supreme Court level to do so (Segal and Spaeth 2001; Rachlinski and Wistrich 2017: 205-206). Rather, these issues are surfaced at the outset to help ensure a fuller understanding of the context of the results of the quantitative analysis. As with gender, conclusions about the influence of judicial attitude given this complexity are suggestive and justify further examination in a broader population.

Magistrate judges were treated differently than district judges in several ways. They cannot take senior status, so this variable was omitted. Race was not clearly identified for most of them directly, so this variable was also omitted. Because they are selected by the sitting judges of the court, it could be asserted that they should reflect the court's median partisanship score. However, in most districts magistrates are nominated, screened and largely retained based on the views of the members of the local bar. (District of Oregon 2017). They are on four or eight-year appointments and a perception of partisanship is grounds for non-renewal. For this reason, it was decided not to calculate scores for magistrates. Instead they were used as a comparison group to judges appointed by Democrats and Republicans in the initial investigation of the impact of partisanship. Cases involving magistrates were also excluded from regression analysis, reducing the maximum n from 510 to 429 rulings.

Of these variables, all but the judicial common space score are nominal. Whether the case was published in the official case reporter (the Federal Supplement), gender, minority status, senior status and whether the judge was appointed before *Kumho* were all coded as dummy variables. Circuit of origin was originally coded to correlate to the numbered circuit solely as a label, then recoded into multiple dummy variables for using this as a variable in certain cross-tabulations and regression. The judge's immediate prior job to appointment was coded to differentiate between private practice, government service and judicial service, then recoded as a dummy variable of prior judicial service for statistical analysis. Table 6 summarizes the details of these variables and how they relate to the hypothesis being analyzed in this study.

Variable Name	Variable Type	Possible Values	Hypotheses Variable Relates To
Case Circuit	Nominal	4=Fourth Circuit 7=Seventh Circuit 9=Ninth Circuit; Recoded to dummy variables for each circuit 1=Fourth 0=Other Circuits; 1=Seventh 0=Other Circuits; 1=Ninth 0=Other Circuits (Two dummies used in each regression variation)	H1-H11
F.Supp.	Nominal/Dummy	0 = Unpublished 1 = Published	Control Only
Judicial Common Space Score	Interval	0.01-1= Conservative Judicial Attitude (larger number =more conservative) -1-(-0.01)=Liberal	H4-H5, H7

		Judicial Attitude (smaller number = more liberal) Recoded to dummy for cross-tabulation as 0=Republican 1= Democrat	
Judge Gender	Nominal/Dummy	0=Male 1=Female	H6
Judge Minority	Nominal/Dummy	0=Non Hispanic White 1=Any other Race/Ethnicity	Control Only
Judge Pre-Kumho	Nominal/Dummy	0=Appointed After 1998 1=Appointed Before 1998	H11
Senior Status	Nominal/Dummy	0= No 1=Yes	H7
Prior Judicial Service	Nominal/Dummy	0= No Prior Judicial Service 1=Prior Judicial Service)	Control Only

Table 6: Court and Judicial Trait Variables.

The third group of variables are those that relate to the specific challenges raised and if they were successful. As noted above, qualitative analysis suggested that judges tend to broadly consider three categories of objections relating to an expert. These are challenges to the expert's credentials, methods and relevance. Some experts are challenged on all dimensions, some only one. Additionally, parties will also often raise a challenge to the propriety of the content of the expert's testimony. This is most commonly a claim that the witness is invading the province of the jury (by commenting on the credibility of other witnesses) or the judge (by providing legal testimony), but it

can also refer to a witness having been improperly coached by counsel or having engaged in improper conduct in their investigation.

These variables are potentially significant for a variety of reasons. Initially, if the treatment of different categories of objections is consistent among all judges regardless of behavioral or attitudinal cues, it suggests that the legal model is being followed consistent with the first hypothesis. Moreover, circuit variation that is consistent with unique precedent would not disprove the value of the legal model but partially validate it, consistent with the second and third hypotheses. This is a distinct question from whether the expert is admitted or not, which has been the focus of most prior studies that may have overlooked this evidence.

Additionally, the possibility that judges are more likely to entertain some objections than others would provide evidence for the idea that judges are engaged in cognitive simplification. Ruling on relevance and the proper scope of testimony are questions that judges rule on in every case and rooted in well-established legal principles all lawyers know, whereas assessing methodology is unique to *Daubert/Kumho* proceedings. Thus, judges are likely far more confident in their ability to understand and process the former rather than the latter. In circumstances like this, Simon (1955: 101-105), Kahneman (2011: 11-12) and Cooper and Neuhaus (2000: 168-171) all suggest people will tend to take action where they feel confident in their expertise and tend to avoid more complicated, cognitively costly tasks. Such a discovery would be consistent with findings in prior studies such as Dixon and Gill (2002: 285-288, 299-300) that judges tend to focus on general acceptance when they do consider the specific factors

mentioned in *Daubert* instead of looking into all the factors equally, and Gatowski, et al.'s (2001: 444-448) finding that judges tend not to understand the concepts involved in assessing *Daubert's* error rate or testability elements.

Finally, distinct from the issue of whether particular objections are granted or not the relationship of the presence of a type of objection and the outcome of the case is an important control in assessing the impact of judicial attitudes. The legal model hypothesizes that legal considerations should be the driving force behind decisions and explain much if not all of the variation between judges. If the judicial attitude model is correct, however, partisanship should be linked to outcomes regardless of the nature of the objections raised in a case. Thus, the actual presence of the objection, in addition to the success of particular types of objections, was coded.

Table 7 summarizes the details of these variables and how they relate to the hypotheses being analyzed in this study.

Variable Name	Variable Type	Possible Values	Hypotheses Variable Relates To
Challenge to Qualifications	Nominal/Dummy	0=No Challenge 1=Challenge Made	H1-H11
Challenge to Qualifications Success	Nominal/Dummy	0=Unsuccessful 1=Successful	H1-H3, H10
Challenge to Methods	Nominal/Dummy	0=No Challenge 1=Challenge Made	H1-H11
Challenge to Methods Success	Nominal/Dummy	0=Unsuccessful 1=Successful	H1-H3, H10
Challenge to Relevance	Nominal/Dummy	0=No Challenge 1=Challenge Made	H1-H11
Challenge to Relevance Success	Nominal/Dummy	0=Unsuccessful 1=Successful	H1-H3, H10
Challenge to Role of Witness/Propriety	Nominal/Dummy	0=No Challenge 1=Challenge Made	H1-H11

Challenge to Role of Witness/Propriety Success	Nominal/Dummy	0=Unsuccessful 1=Successful	H1-H3, H10
--	---------------	--------------------------------	------------

Table 7: Challenge Type and Success Variables.

The last category of variables is the primary dependent variable of interest – the actual outcome of the challenge in terms of what happens to the expert. Although there has to be a successful challenge to have any action taken by the judge, not all challenges are treated the same way. Although prior studies did not mention this, the qualitative analysis revealed the primary response of judges appears to be to restrict rather than exclude the expert. Although restrictions can vary in terms of their significance, they all impact the ensuing trial. Failing to notice this may be part of the reason Helland and Klick (2012: 32-33) concluded *Daubert/Kumho* had no impact on judicial behavior. It may also be that Buchman (2007), who only assessed outcomes, may have coded “grants in part” of motions to exclude as if the expert was excluded rather than restricted yet still testifying.

Given this uncertainty, it was determined to assess these variables both individually and then as a combined “judicial action variable.” In the first scenario, and consistent with what prior studies such as Buchman (2007) and Merlino, Springer and Sigillo (2011) have stated they are analyzing, anything short of a full exclusion is considered an admission and is the dependent variable of interest. In the second, a restriction that did not result in the exclusion of the expert is the dependent variable. This scenario restricts the number of cases under examination somewhat, as a restriction is impossible where an expert is excluded entirely, but as exclusions are relatively rare the n remains in excess of 300 rulings in all variations. Finally, in third scenario, all “judicial

action” is coded as a single act and assessed to determine what influences may make either exclusion or restriction more likely.

Although it might have been simpler to only test exclusions and any judicial action, doing so may have concealed significant findings. The qualitative analysis suggested that arguments about the propriety of testimony – such as that a witness was supplying a legal opinion or usurping the role of the jury – were generally resolved through restriction rather than exclusion of the expert. By comparison, gross failures of methodology, though rare, were more likely to support exclusion as a distinct remedy. Therefore, it was determined to examine each type of action in isolation so as to be able to meaningfully capture both the similarities and differences of the two types of judicial action.

Table 8 summarizes the details of these variables and how they relate to the hypotheses being analyzed in this study.

Variable Name	Variable Type	Possible Values	Hypotheses Variable Relates To
Expert Excluded	Nominal/Dummy	0=Expert admitted 1=Expert excluded	H1-H11
Expert Restricted	Nominal/Dummy	0=Expert not restricted 1=Expert restricted	H1-H11
Judicial Action	Nominal/Dummy	0=No action taken 1=Expert either restricted or excluded	H1-H11

Table 8: Challenge Outcome/Judicial Conduct Variables.

The quantitative analysis consisted of several steps. Initially, frequencies regarding the rate of admissibility, in terms of admissions, restrictions and exclusions,

were compiled for the sample as a whole. Similar data on particular types of challenge and their success were also created.

As an initial step in investigating most of the hypothesized relationships, cross-tabs were compiled as discussed in sections B and C below. This was used to assess the impact of legal factors, partisanship, behavioral traits, strategic and cognitive influences. It was appropriate as a basis for analyzing certain relationships where all relevant variables in question were nominal (Pollock III 2011: 58-61, 212). For example, the third hypothesis - that judges in the 4th Circuit were less likely to use qualifications or relevance as a grounds for rejection owing to the many references to *Kopf* in particular decisions - could be assessed and as discussed below, the null hypothesis could not be rejected. It was also used to determine if the elements hypothesized to influence judicial decision making under the various models had significant relationship with case outcomes on their own prior to inclusion the regression model.⁸ .

B. The Quantitative Evidence is Broadly Consistent with the Qualitative Investigation Regarding the Influence of Law and Cognitive Shortcuts

As discussed in Chapter 3, the central themes emerging from the qualitative investigation were largely consistent with the expectations of the legal model as well as cognitive shortcuts. It appeared that all the judges under review were admitting experts as a default, consistent with the statements in the *Daubert/Kumho* and the Advisory Committee notes to Rule 702 that the exclusion of an expert should be the exception and

⁸ This step also served as a means to assess the behavior of magistrate judges to determine whether they could be meaningfully used as a control for judicial conduct. Despite initial results suggesting magistrates fall in the middle of the partisan spectrum, this proved to be illusory once controls for circuit of origin were added. This led to the decision to exclude magistrates as a class from the ultimate regression model.

admission the default. While they appeared quite willing to rein in testimony by restricting its scope, they were very reluctant to eliminate experts entirely. Judges also generally deferred to other authorities rather than truly investigating the reliability of proposed methods, and viewed most technical objections as matters that went to the weight of the testimony rather than its admissibility.

The first step of the quantitative analysis was designed to confirm the impression that admission was the rule as well as to establish a baseline for examining the first, second, third and sixth hypotheses. To do this, frequencies were established regarding the rate of admissibility, in terms of unconditional admissions, admissions with restrictions and exclusions for all decisions. These results are reported in Table 9. Similar data on the frequency and success rates of particular types of challenge were also created as an element of testing both the tenth and eleventh hypotheses. These results are reported at Table 10. A set of cross-tabs were created that compared the correlation of success on various types of objections with whether an expert was excluded or merely restricted. Those results are reported at Table 11.

Expert Admitted Unconditionally	Expert Admitted with Restrictions	Expert Excluded	Total Cases
268 (52.5%)	150 (29.4%)	92 (18%)	510 (100%)

Table 9: Frequency of Admissions and Exclusions

As can be seen Table 9, both qualitative and quantitative data seem to be in accord at the highest level to strongly support the first hypothesis. Judges are overwhelmingly likely to admit an expert rather than exclude them – 82 percent of

experts were admitted over objections to one degree or another.⁹ This follows *Daubert/Kumho*'s and Rule 702's command to make admission the default and exclusion the exception. That in turn supports the argument that the law is the primary influence in these decisions. The sheer size of the discrepancy argues against this being attributable to anything other than a consensus amongst judges of how to resolve such questions. It is in line with the experimental findings of Kahan, et al (2016: 391-397) that judges tend to be driven by legal considerations in making decisions. Although other factors could also impact outcomes, as an initial piece of data this supports the validity of the first hypothesis and the legal model generally.

At the same time, the data also supports the qualitative impression that judges are very active in monitoring expert behavior and restricting it in appropriate circumstances. Although the vast majority of experts are admitted, 29.4 percent have limits placed on their testimony. When this is coupled to the 18 percent that are rejected, this indicates judges imposed some form of oversight in almost half the cases. This runs counter to the suggestion of Helland and Klick (2012) that *Daubert* is no different than *Frye* in its application. Consistent with the findings of Gatowski, et al (2001: 443-448) it appears judges take their role as "gatekeepers" seriously, even if they are not always comfortable with assessing technical questions about scientific methods. Again, this is what the legal model of judicial decision -making would predict, as judges attempt to follow the law as they understand it. It also supports the interpretation espoused in the tenth hypothesis

⁹ As this study cannot account for experts who testify without objection, it likely understates the overall tendency of such testimony to be admitted into evidence under *Daubert/Kumho*.

that judges are more comfortable with restricting rather than excluding experts because it conforms to norms of judicial conduct within the American adversarial process.

Type of Challenge	Qualifications	Methods	Relevance	Proper Scope of Testimony	Total
Number of Challenges (Percentage of Total Cases)	206 (40.4)	385 (75.5)	205 (40.2)	172 (33.7)	968 (100)
Number of Challenges Upheld (Percentage of Challenges in Category)	65 (31.7)	99 (25.8)	88 (42.9)	114 (66.3)	366 (37.8)

Table 10: Frequency and Success Rate by Challenge Type

The results in Table 10 also buttress the finding in the qualitative review that judges are engaged in some degree of cognitive simplification in resolving these disputes. Objections to the methods used by an expert, while the most frequent in the sample and theoretically being the central focus of a *Daubert/Kumho* proceeding, are also the least likely to succeed. Of the 385 such challenges, only 99 or 25.8 percent were upheld. Qualifications were challenged less frequently, but with slightly higher success at 31.7 percent. Relevance was challenged at almost an identical rate as qualifications, yet was upheld substantially more frequently at 42.9 percent. Although objections to the proper scope of testimony were the least frequent challenge, they were upheld an astonishing 66.3 percent of the time.

This is consistent with the view that judges are choosing to minimize the cognitive cost of their rulings, a typical human behavior as discussed by Simon

(1955:101-105) and Kahneman (2011:31-38). Determining whether methods demonstrate adequate rigor is a unique effort specific to *Daubert/Kumho* and differs with every field of expertise. To really understand these issues would take a significant cognitive investment. When the choice architecture of *Daubert/Kumho* is considered, the path of greatest efficiency for judges on methods challenges would be to reject the challenge in most cases. By comparison, relevance and especially the proper role of the various participants are subjects that lay at the heart of every case and a judge's legal training. That in turn should lead to the development of expertise and confidence in making decisions on this type of objection with relatively low investment of effort, and a greater overall rate of upholding such objections.

The results in Table 11 also show that the majority of sustained objections result in modification of the expert's testimony rather than rejection. This generally supports the tenth hypothesis. Objections to the scope of a witness' testimony, for example, are very commonly upheld, but show a very low correlation with exclusion of the expert rather than restriction. It appears these types of problems with a witness' testimony generally do not merit elimination but restriction as a general rule.

Type of Successful Challenge	Qualifications	Methods	Relevance	Proper Scope of Testimony	Total
Expert Restricted (Percentage)	41 (63.1)	38 (38.4)	47 (53.4)	89 (78.1)	215 (58.7)
Expert Excluded (Percentage)	24 (36.9)	61 (61.6)	41 (46.6)	25 (21.9)	151 (41.3)
Total	65 (100)	99 (100)	88 (100)	114 (100)	366 (100)

Table 11: Frequency of Outcome by Successful Challenge Type

This tendency to restrict rather than exclude testimony is not universal, however. As discussed in Chapter 3, every circuit court's iteration of the *Daubert/Kumho* standard discusses the need to ensure that the testimony is reliable in the sense of being based on a valid method and relevant. These two ideas are the core of *Daubert/Kumho*'s concern that inappropriate testimony be shielded from the jury. Sustained objections to methods result in a significantly higher frequency of exclusions relative to other types of objections. Successful objections to the relevance of testimony, the other mandatory prong of a *Daubert/Kumho* analysis, similarly correlates more strongly with exclusion than either potential problems with qualifications or the proper scope of testimony. This is not consistent with the tenth hypothesis rooted in the legal norm that judges should always favor restriction over exclusion to avoid appearing to interfere in the legal process. It is consistent, however, with the finding in the qualitative investigation that judges are trying to meaningfully apply *Daubert/Kumho* as reported by Gatowski, et al (2001: 443-445). That is, the law instructs judges to ensure that unreliable methods and experts as well as irrelevant testimony are kept from the jury, and when they see such a violation they appear willing to overcome their general reluctance to be seen as interfering with the adversarial system and exclude the witness.

At the same time, this should not be interpreted as an indication that simply making an objection based on methods is more likely to lead to exclusion. Initially, the vast majority of methods based objections are rejected – 384 of the 510 rulings had such an objection, but only 99 or approximately 26 percent were successful. Only 61 such objections or approximately 15.8 percent resulted in exclusions. That is comparable to

exclusions associated with challenges to the proper scope of testimony, which occurred in 25 rulings out of 172 with such an objection or 14.5 percent. Both of these fall between the rates of exclusion associated with successful relevance challenges (24/205 or 11.6 percent) and relevance (41/206 or 19.9 percent). As noted in Part E of Chapter 3, when a judge takes the unusual step of excluding an expert they frequently uphold more than one objection, such as finding both a significant problem with the expert’s qualifications and method, or method and relevance. While Table 11 supports the idea that when a judge finds methods to be problematic they are more willing to exclude rather than restrict the testimony, it does not speak to when a judge is likely to find an objection to methods – or any other objection – to be adequately supported as a matter of law.

Although the global data supports the themes identified in the qualitative review, it is still not entirely monolithic. Once the decisions are broken down by circuit of origin, some distinctions become clear. As shown in Table 12 and Figure 4, as predicted by the second hypothesis, judges presiding in districts within the Ninth Circuit are substantially more likely to exclude expert testimony than their colleagues in the Fourth and Seventh Circuits. These differences are statistically significant based on chi-square and the association of location and outcome is robust under Cramer’s V. At a minimum this supports further investigation of the impact of circuit level variation.

Circuit of Origin	4 th Circuit	7 th Circuit	9 th Circuit	Total
Expert Not Excluded (Percentage in Circuit)	78 (84.8)	181 (87.9)	159 (75)	418 (82)
Expert Excluded	14(15.2)	25 (12.1)	53 (25)	92 (18)

(Percentage in Circuit)				
Total (Percentage in Circuit)	92 (100)	206 (100)	212(100)	510 (100)
Pearson's Chi-Square/P-value				12.298/.002
Lambda				.000
Cramer's V/P-value				.155/.002

Table 12: Exclusion of Challenged Experts by Circuit

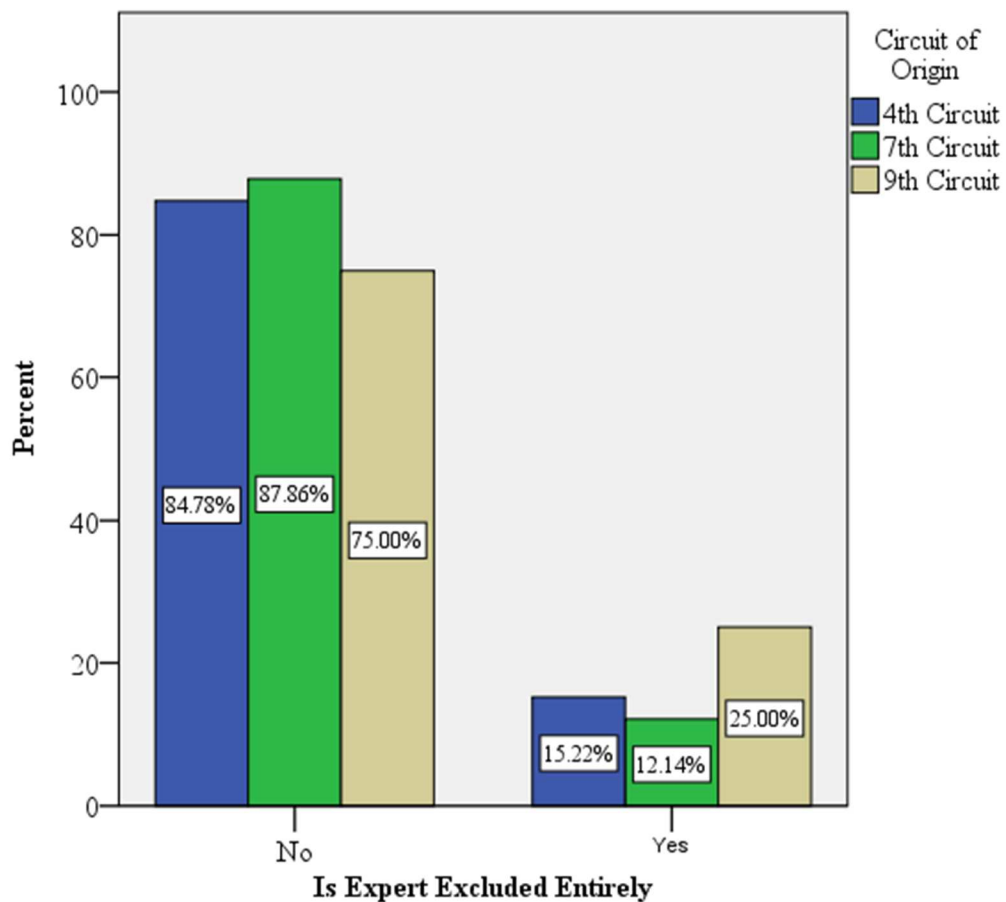


Figure 4: Rates of Admission and Exclusion by Circuit

While this is at odds with the general reputation of the Ninth Circuit as a liberal bastion, it is consistent with the Ninth Circuit's precedent in *Mukhtar*. Judges in the Ninth Circuit, laboring under the risk of having to conduct a new trial if they are found to

have made insufficient findings rather than just making some new findings to justify their action, should be more careful. This is also consistent with both the legal model and the cognitive influence of choice architecture. At the same time, the overall admission rate is still 75 percent, consistent with *Daubert/Kumho*'s and Rule 702's express preferences and the first hypothesis that judges will default to admitting experts. Location does appear to matter, however, and suggests that studies of judicial decision making should look at these circuit level variations rather than assume circuits are interchangeable.

Moreover, local disparities extend beyond the act of exclusion. As documented in Table 13 and Figure 5, judges within each circuit were generally more likely to restrict testimony rather than exclude a witness. However, the Ninth Circuit still showed a relatively higher frequency of restricting testimony in cases where the expert is not excluded. The restriction rate in the Ninth Circuit is 44 percent, substantially higher than the other two locations examined. Intriguingly, there is also a split between the Fourth and Seventh Circuit, with the Seventh Circuit acting in only 29 percent of cases while the Fourth Circuit restricted testimony in nearly 35 percent. These relationships were significant but slightly more attenuated than in the case of restriction based on the reduction in the Cramer's V score to .139.

Circuit of Origin	4 th Circuit	7 th Circuit	9 th Circuit	Total
Expert Not Restricted (Percentage in Circuit)	51 (65.4)	128 (70.7)	89 (56.0)	268 (64.1)
Expert Restricted (Percentage in Circuit)	27 (34.6)	53 (29.3)	70 (44)	150 (35.9)

Total (Percentage in Circuit)	78 (100)	181 (100)	159 (100)	418 (100)
Pearson's Chi-Square/P-value				8.064/.018
Lambda				.000
Cramer's V/P-value				.139/.018

Table 13: Restrictions of Expert Testimony by Circuit in Cases Where Expert Not Excluded

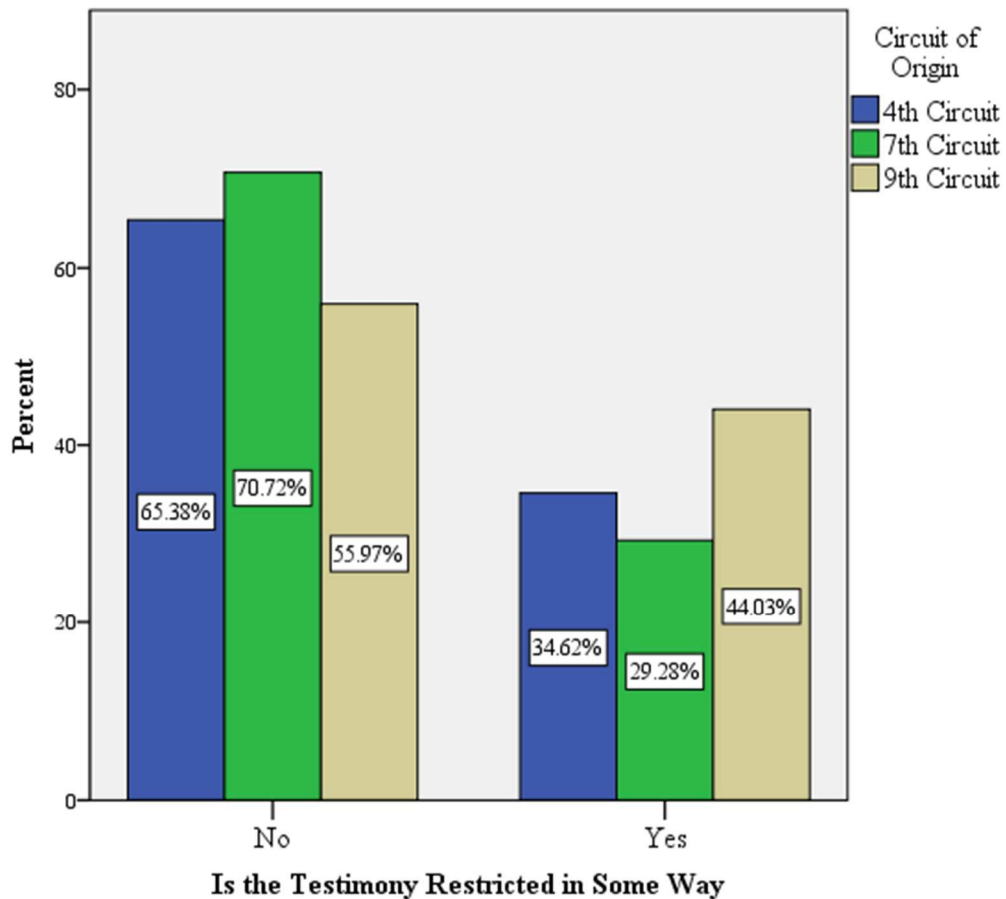


Figure 5: Restrictions of Expert Testimony by Circuit in Cases Where Expert Not Excluded

Combining exclusions and restrictions into a single judicial action variable in

Table 14 underscores the degree with which judges are actively applying *Daubert/Kumho* as first reported by Gatowski et al (2001). While they may not be eager to exclude

witnesses, judges appear to be quite willing to take some action where they deem it appropriate. The lowest rate of action was approximately 38 percent in the Seventh Circuit, suggesting a significant degree of scrutiny. At the same time, differences linked to location continued to be significant, demonstrating a moderate relationship between location and the likelihood of action. Even so, the Ninth Circuit still stands out in the degree to which judges seem willing to take action. Whereas the default overall is to admit an expert, judges find a basis to take action in response to 58 percent of all challenges. While they are honoring the stated goals of *Daubert/Kumho* and Rule 702 to admit most witnesses, there is a robust degree of gatekeeping in action, in contrast to findings in studies like Helland and Klick (2011) which focused solely on admission rates.

Circuit of Origin	4 th Circuit	7 th Circuit	9 th Circuit	Total
No Judicial Action (Percentage in Circuit)	51 (55.4)	128 (62.1)	89 (42.0)	268 (52.5)
Expert Restricted or Excluded (Percentage in Circuit)	41 (44.6)	78 (37.9)	123 (58)	242 (47.5)
Total (Percentage in Circuit)	92 (100)	206 (100)	159 (100)	510 (100)
Pearson's Chi-Square/P-value				17.395/.000
Lambda				.140
Cramer's V/P-value				.185/.000

Table 14: Overall Rates of Judicial Action by Circuit

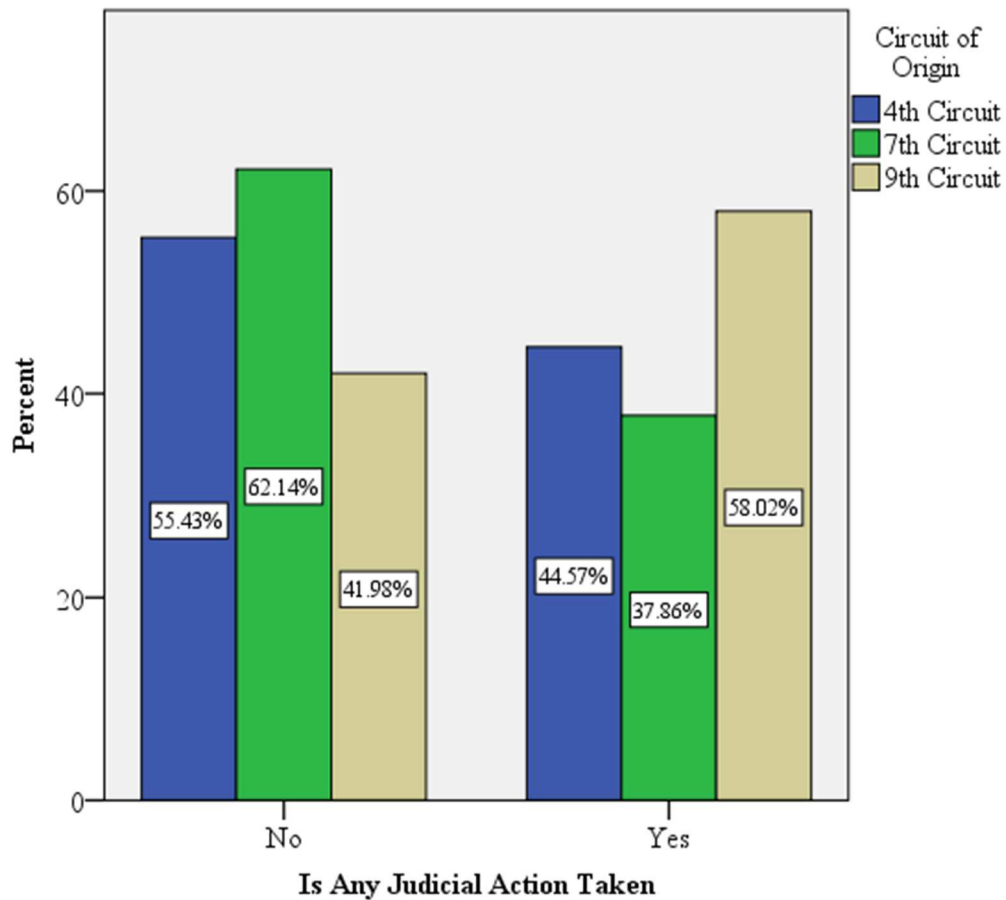


Figure 6: Overall Rates of Judicial Action by Circuit

Although this speaks to the possibility of significant inter-circuit distinctions that are still consistent with the legal model of decision making, not all hypothesized relationships were borne out by the data. In particular, the suggestion that the repeated citation of *Kopf* by judges in the Fourth Circuit was meant to signal hostility to challenges based on the qualifications or relevance of the expert was not borne out by the data. As shown in tables 15 and 16, although these judges had substantially fewer rulings overall, and as shown previously in table 8, admitted over 80 percent of all experts, they demonstrated a significantly greater propensity to uphold objections to qualifications or relevance than judges in either the Seventh or Ninth Circuit. It may be that local practice

and norms, including some internalization of *Kopf*, tends to discourage all but the most convincing motions to restrict or exclude from being brought. Ultimately, though, the language of *Kopf* does not carry over to how judges evaluate the disputes before them. For this reason, the null alternative to the third hypothesis could not be rejected and this relationship was not analyzed further in this study.

Circuit of Origin	4 th Circuit	7 th Circuit	9 th Circuit	Total
Challenges to Qualifications Rejected (percentage in circuit)	19 (46.3)	75 (80.6)	46 (64.8)	89 (42)
Challenges to Qualifications Upheld (percentage in circuit)	22 (53.7)	101(19.4)	25 (35.2)	123 (58)
Total	41 (100)	93 (100)	71 (100)	205 (100)
Pearson's Chi-Square/P-value				16.079/.000
Lambda				.046
Cramer's V/P-value				.280/.000

Table 15: Rates of Upholding Challenges to Qualifications by Circuit

Circuit of Origin	4 th Circuit	7 th Circuit	9 th Circuit	Total
Challenges to Relevance Rejected (percentage in circuit)	15 (45.5)	62 (67.4)	40 (50)	117 (57.1)
Challenges to Relevance Upheld (percentage in circuit)	18 (54.5)	30 (32.6)	40 (50)	88 (42.9)
Total	33 (100)	92 (100)	80 (100)	205 (100)
Pearson's Chi-Square/P-value				7.450/.024
Lambda				.034
Cramer's V/P-value				.191/.024

Table 16: Rates of Upholding Challenges to Relevance by Circuit

It appears that the themes of the qualitative investigation are largely confirmed by the quantitative data. Judges in every court appear to be following the broad precepts of *Daubert/Kumho* and admitting most witnesses. At the same time they are engaging in a

significant degree of oversight including restrictions on testimony, and there are some significant distinctions between different circuits in terms of which objections are upheld, rates of exclusion and overall judicial action. Their behavior is also broadly consistent with cognitive influences such as following choice architecture, minimizing cognitive burdens and avoiding cognitively costly tasks like reviewing technical information and methods in depth. Although the evidence of cognitive influences is more suggestive and the evidence of adherence to legal precepts more conclusive, the first, second, ninth and tenth hypothesis are all consistent with the evidence in the dataset.

C. There is Suggestive Evidence for Multiple Influences on Judicial Decision Making

Although the initial quantitative analysis is broadly consistent with the themes of the influence of both the legal and cognitive models of decision making identified in the qualitative investigation, this does not exclude the potential influence of other factors. It could be that the apparent consensus is an artificial mean between wildly opposed partisans. Any number of behavioral traits or strategic considerations may also be present, but their relative infrequency causes the results to be masked by the dominant patterns of behavior. Testing the potential influence of partisanship, behavioral traits and strategic considerations – items which could not be assessed in the qualitative review but are suggested by the literature and caselaw on judicial decision making – reveals a number of potential additional factors that may be influencing judicial decision making in this context. In almost every case, though, they are of limited value and reveal that there is no single factor that compares to the dominant consensus of judges in favor of admitting expert testimony.

To create an initial assessment of the role of partisanship and test the fourth hypothesis, the frequencies compiled in Part B of this chapter for exclusions, restrictions and the overall rate of judicial action across the entire sample were re-analyzed in a series of cross-tabs that controlled for partisanship and were limited to Article III judges. The findings of this set of analyses are reported in Tables 17, 18 and 19.

Judge Alignment	Democratic Appointee	Republican Appointee	Total
Experts Admitted (percent)	172 (85.1)	179 (78.9)	351 (81.8)
Experts Excluded (percent)	30 (14.9)	48 (21.1)	78 (18.2)
Total	202 (100)	227 (100)	429(100)
Person Chi-Square/P-value			2.846/.092
Lambda			.000
Cramer'sV/P-value			.081/.092

Table 17: Relationship of Partisanship to Exclusion of Experts

Judge Alignment	Democratic Appointee	Republican Appointee	Total
Experts Admitted Without Restriction (percent)	118 (68.6)	107 (59.8)	225 (64.1)
Experts Admitted with Restrictions (percent)	54 (31.4)	72 (40.2)	126 (35.9)
Total	172	179	351
Person Chi-Square/P-value			2.971/.085
Lambda			.037
Cramer'sV/P-value			.092/.085

Table 18: Relationship of Partisanship to Restriction of Experts

Judge Alignment	Democratic Appointee	Republican Appointee	Total
No Judicial Action (percent)	118 (58.4)	107 (47.1)	225 (52.4)
Experts Restricted or Excluded	84 (41.6)	120 (52.9)	204 (47.6)

(percent)			
Total	202 (100)	227 (100)	429 (100)
Person Chi-Square/P-value			5.452/.020
Lambda			.059
Cramer's V/P-value			.113/.020

Table 19: Relationship of Partisanship to Overall Judicial Action

As shown in Tables 17 and 18, there is not a robust difference between judges at a national level in regards to exclusion or restrictions based on partisanship. Judges appointed by Democratic Presidents are, consistent with Buchman (2007), less likely to exclude or restrict expert testimony than judges appointed by Republican Presidents. At the same time, the overall rate of disagreement is actually smaller than what Kahan, et al (2016: 391-397) found experimentally in a study that concluded judges were generally insulated from the impacts of personal values in deciding cases. It also is smaller than the degree of variation exhibited by experts in a number of fields like art, wine or medicine where politics or policy preferences are not generally thought to impact outcomes (Galanter 1993: 90-91; Baker 2005: 502-503). Furthermore, the relationship in both cases lacks statistical significance. This indicates the null hypothesis – that partisanship does not impact these decisions - cannot be rejected at this time. However, in light of the localized distinctions found in Part B as well as the different results obtained by Buchman (2007: 680-681) when using more granular measures of partisanship in regression analysis, further refinement and testing is likely justified before rejecting the value of partisanship entirely

However, when the combined judicial action variable is used a different result appears. In combination the gap between Republicans and Democrats expands to 11.3

percent. While still smaller than the differences reported by Kahan et al (2016), Galanter (1993) and Baker (2005) in various studies of judgment, it does achieve statistical significance at the .05 level. This suggests that partisanship does matter, although it is marginal relative to the overall consistency of judicial action.

To assess if this was truly a consistent phenomenon, the results were broken down by circuit while controlling for partisanship. The exceptionally low number of rulings in the Fourth Circuit by Democratic appointees suggested partisanship could not be assessed within that circuit. For this reason, controlled comparisons on exclusions, restrictions and judicial action excluding magistrates and limited to the Seventh and Ninth Circuits were conducted. These results are shown in Tables 20, 21 and 22 below.

	Judge Alignment	Democratic Appointee	Republican Appointee	Total
7 th Circuit	Experts Admitted (percent)	78 (91.8)	62 (83.8)	140 (88.1)
	Experts Excluded (percent)	7 (8.2)	12 (16.2)	19 (11.9)
	Total	85 (100)	74 (100)	159 (100)
	Person Chi-Square/P-value			2.395/.122
	Lambda			.000
	Cramer's V/P-value			.123/.122
9 th Circuit	Experts Admitted (percent)	90 (79.6)	63 (73.3)	153 (76.9)
	Experts Excluded (percent)	23 (20.4)	23 (26.7)	46 (23.1)
	Total	113 (100)	86 (100)	199 (100)
	Person Chi-Square/P-value			1.122/.289
	Lambda			.000
	Cramer's V/P-value			.075/.289
Total	Experts Admitted (percent)	168 (58.4)	125 (47.1)	293 (52.5)
	Experts Excluded (percent)	30 (41.6)	35 (52.9)	65 (47.5)

	Total	198(100)	160 (100)	358 (100)
	Person Chi-Square/P-value			2.692/.101
	Lambda			.000
	Cramer'sV/P-value			.087/.101

Table 20: Relationship of Partisanship to Exclusion of Experts Controlling for Location, Limited to cases in Seventh and Ninth Circuits

	Judge Alignment	Democratic Appointee	Republican Appointee	Total
7 th Circuit	Experts Admitted Without Restriction (percent)	65 (83.3)	38 (61.3)	103 (73.6)
	Experts Restricted (percent)	13 (16.7)	24 (38.7)	37 (26.4)
	Total	78 (100)	62 (100)	140 (100)
	Person Chi-Square/P-value			8.632/.003
	Lambda			.000
	Cramer'sV/P-value			.248/.003
9 th Circuit	Experts Admitted Without Restriction (percent)	49 (54.4)	37 (58.7)	86 (56.2)
	Experts Restricted (percent)	41 (45.6)	26 (41.3)	67 (43.8)
	Total	90 (100)	63 (100)	153 (100)
	Person Chi-Square/P-value			.277/.599
	Lambda			.000
	Cramer'sV/P-value			.043/.599
Total	Experts Admitted Without Restriction (percent)	114 (67.9)	75 (60)	189 (64.5)
	Experts Restricted (percent)	54 (32.1)	50 (40)	104 (35.5)
	Total	168(100)	125 (100)	293 (100)
	Person Chi-Square/P-value			1.933/.164
	Lambda			.000
	Cramer'sV/P-value			.081/.164

Table 21: Relationship of Partisanship to Restriction of Experts Controlling for Location, Limited to cases in Seventh and Ninth Circuits

	Judge Alignment	Democratic Appointee	Republican Appointee	Total
7 th Circuit	No Judicial Action (percent)	65 (76.5)	38 (51.4)	103 (64.8)
	Judicial Action (percent)	20 (23.5)	36 (48.6)	56 (35.2)
	Total	85 (100)	74 (100)	159 (100)
	Person Chi-Square/P-value			10.940/.001
	Lambda			.000
	Cramer's V/P-value			.262/.001
9 th Circuit	No Judicial Action (percent)	49 (43.4)	37 (43)	86 (43.2)
	Judicial Action (percent)	64 (56.6)	49 (57)	113 (56.8)
	Total	113 (100)	86 (100)	199 (100)
	Person Chi-Square/P-value			.002/.962
	Lambda			.000
	Cramer's V/P-value			.003/.962
Total	No Judicial Action (percent)	114 (57.6)	75 (46.9)	189 (52.8)
	Judicial Action (percent)	84 (42.4)	85 (53.1)	169 (47.2)
	Total	198 (100)	160 (100)	358 (100)
	Person Chi-Square/P-value			4.066/.044
	Lambda			.059
	Cramer's V/P-value			.107/.044

Table 22: Relationship of Partisanship to Judicial Action Controlling for Location, Limited to cases in Seventh and Ninth Circuits

Partisanship seems to matter quite a bit in the Seventh Circuit. The between Seventh Circuit judges appointed by Democrats and those appointed by Republicans is more pronounced with regards to restrictions and overall action than it is with regards to exclusion. In the context of restrictions, where the law does not indicate a default position to take, the split between the two groups in the Seventh Circuit swelled from a

relatively modest 8 point gap to 22 points, nearly tripling in size. The relationship of partisanship to action is statistically significant at the .005 level, and the Cramer's V score suggests a moderately strong relationship. This pattern persisted when cases containing restrictions or exclusions were combined in the judicial action variable. The relationship in the Seventh Circuit was strong enough to make the relationship appear significant in the entire sample.

At the same time, the Ninth Circuit showed a fundamentally different pattern. Although Democratic appointees were slightly less likely to exclude testimony than their Republican appointed counterparts, they were actually more likely to restrict testimony and the overall propensity to take action was almost identical. Democratic appointees also consistently demonstrated a greater propensity to act than either group of judges in the Seventh Circuit. This suggests supporting the second hypothesis – judges in the Ninth Circuit are more inclined to act than others owing to local authority. At the same time, the results in the Seventh Circuit suggest that a revised version of the fourth hypothesis can also be supported – in the absence of controlling law that shifts judicial actions, partisanship can play a significant role in the actions of judges, with judges appointed by Republicans more likely to exclude or restrict experts than those appointed by Democrats.

Buchman (2007:687) also finds that challenges brought by corporate parties against individuals are more successful generally, and suggests that conservative judges should be more inclined to restrict experts. At one level this makes intrinsic sense – if liberal judges are concerned with the individual or one-shot players in the law, where

conservatives favor corporate and other repeat players in the language of Galanter (1974: 96-103), then it would follow that a challenge by a corporate defendant to the testimony of an individual plaintiff's expert should lead to the greatest divide between judges with different partisan outlooks. When these cases were actually considered, however, the results conflicted with expectations. As shown in Table 23, when considering the relationship of these factors to restrictions, partisanship was only statistically significant outside of the context that was supposed to favor partisanship – the gap between the groups almost disappeared in the context of repeat players challenging individuals. On exclusions, the results were even stranger, as Democratic appointees actually increased their likelihood of excluding experts retained by individuals whereas Republican appointees were slightly less likely to exclude an individual's expert when a corporate or government defendant brought the challenge.

	Judicial Alignment	Democrat	Republican	Total
Corp. or Govt Def. Challenge + Ind. Plaintiff	Experts Admitted Without Restriction (percent)	69 (69)	64 (66.7)	133 (67.9)
	Experts Restricted (percent)	31 (31)	32 (33.3)	63 (32.1)
	Total	100 (100)	96 (100)	196 (100)
All other challenges	Expert Admitted Without Restriction (percent)	49 (68.1)	43 (51.8)	92 (59.4)
	Experts Restricted (percent)	23 (31.9)	40 (48.2)	63 (40.6)
	Total	72 (100)	83 (100)	155 (100)
Chi-square/P- value				2.971/.085
Lambda				.000
Cramer's V/P- Value				.092/.085
Corp. or Govt Def. Challenge	Experts Admitted (percent)	100 (82.6)	96 (80)	196 (81.3)

+ Ind. Plaintiff				
	Experts Excluded (percent)	21 (17.4)	24 (20)	45 (18.7)
	Total	121 (100)	120 (100)	241 (100)
All other challenges	Experts Admitted (percent)	72 (88.8)	83 (77.6)	155 (82.4)
	Experts Excluded (percent)	9 (11.2)	24 (22.4)	33 (17.6)
	Total	81 (100)	107 (100)	188 (100)
Chi-square/P-value				2.846/.092
Lambda				.000
Cramer's V/P-Value				.081/.092

Table 23: Impact of Partisanship on Restriction and Exclusion of Experts Controlling for Nature of Party Bringing Challenge

Breaking this element down by circuit again pointed to the significance of location but did not support the fifth hypothesis.¹⁰ The Seventh Circuit follows the prediction – judges appointed by Democrats are less likely to exclude or restrict a witness in this context, whereas the restriction and exclusion rate for those appointed by Republicans goes up. Even so, it was not a statistically significant relationship in any iteration of the dependent variable within this circuit or overall. In fact, the only place this variable could be substantially linked to partisanship was in the Ninth Circuit, and that was in the opposite direction than expected. That is, the rate of restrictions by judges appointed by Democratic Presidents went up in these circumstances, not down – and actually exceeded the restriction rate of judges appointed by Republican Presidents. Again, this supports the conclusion that the null alternative to the fifth hypothesis cannot

¹⁰ In light of these results, the representation of this data in a table would needlessly occupy excessive space and was therefore omitted. The SPSS output documenting these results is available from the author.

be rejected in this particular dataset, while underscoring the significance of understanding circuit level differences.

At this level, partisanship's influence is ambiguous. There is a consistent divide between Democrat and Republican appointees within different circuits, but its impact is not always consistent with expectations. One further question though is whether the partisanship gap reflects sincere belief or attempts at “auditioning” for elevation to the next level of the judiciary. As Posner (2008: 159) notes, taking senior status eliminates any incentive to try and prove party loyalty in the hopes of being promoted to the appellate level. Additionally, senior status judges can be barred from hearing cases if they act too aggressively. To assess that, cross tabulations were constructed that compared rates of exclusions and restrictions between active and senior status judges generally, then controlled for partisanship.

Ultimately, senior status had almost no impact on exclusion rate. Active judges excluded 18.4 percent of cases, and senior judges excluded 19.3 percent. There was a slightly larger gap on restrictions – active status judges restricted experts in 36.6 percent of cases, senior judges in 30.4 percent – but this was not statistically significant. Introducing a control for partisanship suggested that political differences, while not statistically significant, were persistent regardless of status. Democratic appointees were less likely to exclude or restrict experts whether they were active or on senior status. Republican judges on senior status did restrict at a lower rate than their active compatriots – 32.3 versus 41.9 percent – but this was not a statistically significant difference. Based on this finding, the null hypothesis to the seventh hypothesis – which

asserted judges on senior status would be less partisan than active status judges – cannot be rejected.

In addition to assessing the role of partisanship and location, the literature suggests that several behavioral variables might matter to the outcome. Specifically, the gender and the time of appointment relative to the completion of the *Daubert/Kumho* trilogy theoretically might have an appreciable impact. Boyd (2013: 211-213) suggests that women are socialized to solve problems and as judges may be less attached to the adversarial model rather than finding ways to resolve cases. This might extend to a greater willingness to restrict or exclude experts as a route to creating a basis for settlement. A similar but opposite impact might come from judges appointed before *Daubert/Kumho*. If they had ingrained the older attitude of non-interference on issues relating to expert testimony found by Cecil and Willging(1993: 18-21) these judges might be less likely to act.

Each of these factors was tested against the rates of exclusion and restriction. Given the relatively small population of rulings made by judges fitting into any of these categories, controls for location and partisanship were not initially included. The rates of restriction for male and female judges were almost identical - 36.3 percent for the former and 34.2 percent for the latter. On exclusions, women were slightly more willing to act than men – 21.6 percent versus 17.2 percent – but it was not statistically significant. Based on these results, it appears that in the case of the sixth hypothesis, suggesting that female judges would be more likely to act than male judges, the null hypothesis cannot be rejected.

The impact of a judge's time of appointment presented a more complex question. On exclusions, the rates were almost identical – judges appointed after the *Daubert/Kumho* trilogy was completed excluded 18 percent of all experts, whereas those appointed before then excluded 18.2 percent of cases. On restrictions, however, there was a significant difference – the more recently appointed judges restricted testimony in 38.4 percent of cases whereas those appointed earlier only restricted testimony in 23.6 percent of cases. At a minimum this suggests that the eleventh hypothesis cannot be rejected out of hand.

Controlling for partisanship, reported in Table 24, revealed that the distinction was primarily associated with one group of judges. Almost 70 percent of rulings by judges appointed before *Kumho* come from judges appointed by Republican Presidents. This cohort showed a marked difference with their fellow Republican appointees appointed post *Kumho*. The latter group restricted testimony in 45.8 percent of cases, whereas the former acted in only 25 percent of the cases. Although this was the only subgroup that showed a statistically significant relationship between time of appointment and conduct, this effect was pronounced enough to make the relationship statistically significant for the entire dataset as well. It is possible that there are other explanations specific to the shift in Republican appointments. However, at a minimum it provides support for the eleventh hypothesis, and suggests a basis for further investigation to assess if this is a unique effect for Republican appointees or would apply more broadly if more Democratic appointees were examined.

	Judge Appointment	Post- <i>Kumho</i>	Pre- <i>Kumho</i>	Total
Magistrates	Experts Admitted	33 (60)	10 (83.3)	43 (64.2)

	without Restrictions (percent)			
	Experts Restricted (percent)	22 (40)	2 (16.7)	24 (35.8)
	Total	55 (100)	12 (100)	67 (100)
	Person Chi-Square/P-value			2.333/.127
	Lambda			.000
	Cramer's V/P-value			.187/.127
Democratic Appointees	Experts Admitted without Restrictions (percent)	109 (68.1)	9 (75)	118 (68.6)
	Experts Restricted (percent)	51 (31.9)	23 (26.7)	54 (31.4)
	Total	160 (100)	12 (100)	172 (100)
	Person Chi-Square/P-value			.245/.621
	Lambda			.000
	Cramer's V/P-value			.038/.621
Republican Appointees	Experts Admitted without Restrictions (percent)	71 (54.2)	36 (75)	107 (59.8)
	Experts Restricted (percent)	60 (45.8)	12 (25)	72 (40.2)
	Total	131 (100)	48 (100)	179 (100)
	Person Chi-Square/P-value			6.322/.012
	Lambda			.000
	Cramer's V/P-value			.188/.012
Total	Experts Admitted without Restrictions (percent)	213 (61.6)	55 (76.4)	268 (64.1)
	Experts Restricted (percent)	133 (38.4)	17 (23.6)	150 (35.9)
	Total	346 (100)	72 (100)	418 (100)
	Person Chi-Square/P-value			5.696/.017
	Lambda			.000
	Cramer's V/P-value			.117/.017

Table 24: Impact of Appointment Prior to Completion of *Daubert/Kumho*, Controlling for Partisanship

Finding factors that could be used to assess the strategic and institutional models proved somewhat elusive in this dataset. As discussed above, one aspect of the data that speaks to the validity of both a legal and strategic model is the apparent tendency of courts in the Ninth Circuit to be stricter in keeping with local precedent. Another possible marker that occurred with some frequency was that an objection to an expert witness would be made as part of, or consolidated with, a motion for summary judgment. Frequently but not universally a defendant was seeking to exclude an expert and asserting that in doing so the entire case would fail. This in turn would create a strategic incentive for judges to act, as the early elimination of cases is favored both by individual judges and the court system as a whole (Boyd 2013: 197-198; Robbennolt, MacCoun, and Darley 2010: 29). Dixon and Gill (2002: 295-296) noted that *Daubert/Kumho* motions seemed to be associated with a rise in summary judgements based on expert testimony but did not compare the relative rate of exclusion at the summary judgement stage to other contexts.

To test this proposition, two cross-tabs were constructed, one each for restrictions and exclusions of experts. Initially, the high rate of challenges that were involved in summary judgment proceedings (152 out of 510 total challenges, or 29.8 percent) suggests attorneys see value in this approach. As documented in Tables 25 and 26, the results ran in opposite directions. Restrictions go down to a significant degree in the context of a summary judgment motion, from 39.4 to 26.7 percent. This relationship is only moderate based on a Cramer's V value of 0.118. At the same time, exclusions rose

substantially, from 15.6 to 23.7 percent. This relationship is also statistically significant at the .05 level, but there is only a weak relationship based on a Cramer's V value of .096.

This is a reversal of the general trend in the data, in which restrictions are more common than exclusions. However, it makes sense in the context of summary judgment motion. As noted exclusion has a strategic advantage as it can facilitate settlement, narrow issues for trial or terminate a case entirely. Short of exclusion, however, restrictions at the summary judgment stage would lack much value for a judge looking to eliminate a case or spur resolution. Frequently judges put off the final resolution of *Daubert/Kumho* and similar issues for as close to trial as possible, likely due to awareness of the propensity of cases to settle before trial. In fact, judges in this study often declined to rule on a *Daubert/Kumho* issue in a summary judgment motion where they determined a final ruling was not necessary to pass on the summary judgment stage.

Summary Judgment Motion	No	Yes	Total
Experts Admitted Without Restriction (percent)	183 (60.6)	85 (73.3)	268 (64.1)
Experts Restricted (percent)	119 (39.4)	31 (26.7)	150 (35.9)
Total	302 (100)	116 (100)	418 (100)
Pearson's Chi-Square/P-value			5.857/.016
Lambda			.000
Cramer's V/P-value			.118/.016

Table 25: Relationship of Summary Judgment Motion and Restriction of Expert Testimony

Summary Judgment Motion	No	Yes	Total
Experts Admitted	302 (84.4)	116 (76.3)	418 (82)

(percent)			
Experts Excluded (percent)	56 (15.6)	36 (23.7)	92 (18)
Total	358 (100)	152 (100)	510 (100)
Pearson's Chi-Square/P-value			4.667/.031
Lambda			.000
Cramer's V/P-value			.096/.031

Table 26: Relationship of Summary Judgment Motion and Exclusion of Expert Testimony

Given the significance of location and partisanship shown above, further cross tabs were constructed that controlled for each of these factors. An additional set included both variables as different layers of control. In every case the pattern was the same – restrictions went down in the context of summary judgment motions while exclusions went up. However, almost none of the subgroups created by introducing controls were statistically significant.¹¹

Once location was introduced as a control, for example, judges in each region showed consistent results, but only one subgroup out of six demonstrated statistical significance. In the Fourth Circuit, restrictions went down from 36.7 to 31.0 percent while exclusions rose from 12.5 to 19.4 percent. In the Seventh Circuit, restrictions went from 34 to 10.8 percent, while exclusions went from 10 to 19.6 percent. In the Ninth Circuit, the restrictions went from 39.4 to 26.7 percent, while exclusions went from 23.2 to 28.6 percent. However, only the difference on restrictions in the Seventh Circuit was statistically significant at the .05 level.

¹¹ Given the prolix nature of the cross tabulations once these multiple levels were introduced, the representation of this data in a table would needlessly occupy excessive space and was therefore omitted. The SPSS output documenting these results is available from the author.

Even when location and partisanship were included, the pattern persisted in every subgroup – the percentage of restrictions went down, exclusions went up. At the same time few cells within the cross-tabs were shown to be statistically significant, and then only in particular contexts. For example, Republican appointed judges in the 7th Circuit showed a statistically significant increase in exclusions, but the reduction in restrictions for this cohort was not. Conversely the reduction in restrictions by magistrates in the 4th Circuit for restrictions was statistically significant, but not the increase in exclusions.

Ultimately this examination provides a complicated answer about the role of summary judgment motions as a strategic vehicle for judges. There is not a conclusive basis for rejecting the null hypothesis in all cases. At the same time, the uniformity of result, and the particularized findings of significance suggest that conducting the same examination on a larger dataset might provide that evidence. It might also document that the ninth hypothesis that judges are acting strategically at the summary judgment stage of cases is valid only under particular conditions with regard to location, partisanship or other factors.

The ambiguity of the results of the initial examination of the ninth hypothesis is similar to several other factors considered in this study. There seems to be little to no evidence to support the impact of certain factors, and certainly no finding adequate to reject the null hypothesis. Gender, for example, does not appear to have an influence in these decisions. This in turn underscores the finding in Boyd (2016: 793-796) that issues with personal salience to women tend to show a gendered split while more neutral issues such as expert witness admission do not. A judge's senior or active status also does not

seem to impact these outcomes, either. This suggests either that Posner's (2008: 159) contention that judges on senior status are risk averse owing to the conditional nature of their job is inaccurate or that judges do not see these decisions as carrying an especially salient degree of risk. In either event, status does not seem to be related to the outcome of these rulings.

At the same time, even these negative findings have implications for the dominant models in the field of judicial decision making. The fact that judges, regardless of gender, act similarly on questions of admissibility suggests that judicial behavior primarily follows the legal model and judges tend to have similar understandings of the impact of the law as stated in *Daubert/Kumho*. Although the persistence of partisanship past retirement is consistent with the judicial attitude model, the marginal nature of it as a factor and suggests that the model should be used with caution in assessing the conduct of trial court judges.

Of perhaps greater interest is that several factors examined were shown to influence judicial decision making, at least in some circumstances. Partisanship appears to matter, but its influence is dependent on the issue and the circuit in which a case originates. This simultaneously supports Buchman's (2007:680-690) findings while refuting his further contention that the nature of the parties to a case and which party brings the challenge against the expert has any predictable impact, much less a partisan one. The timing of a judge's appointment seems to matter to some degree as well, suggesting a shift in the social conditioning of judges in the years after *Daubert/Kumho* became the national standard for evaluating all expert testimony. Finally, it appears that

judges are influenced by when a challenge is made to an expert's testimony. The rise in the rate of exclusions when the challenge is part of a summary judgment proceeding suggests some degree of strategic consideration, as the rejection of the expert can facilitate the resolution of disputes short of trial. The simultaneous decline in restrictions on testimony in this setting, on the other hand, is consistent with a finding that judges conserve cognitive resources through avoidance of what would be preliminary and potentially pointless findings. As a whole, this portion of the study suggests there are a number of fruitful avenues for further investigation of *Daubert/Kumho* challenges as a means of gaining a more complete understanding of the nature of judicial decision making at the district court level.

D. Regression Analysis Supports the influence of Law, Cognition and Political Affiliation on Judges

As discussed in Parts A-C above, there is good reason to think a multitude of factors influence judicial decision making at the district court level. Some, such as the idiosyncratic nature of individuals, defy quantitative analysis or any other generalized assessments. Others, like the overwhelming tendency of judges to follow the exhortation in *Daubert/Kumho* and the Advisory Committee comments to Rule 702 to admit most experts, are difficult to assess simply because they show such little variation. Still, variables including what types of challenges are brought in a case, the location the case is brought in, partisanship of judges, the timing of a challenge in a life cycle of a case and when a judge was appointed all appear to have some influence on these judges. These various factors suggest that each of the models discussed in Chapter 2 – legal,

behavioral/attitudinal, strategic/institutional and cognitive – may help explain judicial behavior, although the law seems to have more of an effect than any other factor.

Based on the results of the initial investigation, a regression model capturing key variables related to the potential influence of legal, behavioral, attitudinal, strategic, institutional and cognitive influences was constructed. The dependent variable for assessing most of the hypotheses is the outcome of the case, in terms of the status of the expert. As this is a binary concept, whether expressed as restrictions on testimony, the exclusion of experts or the combined sum of judicial actions, bivariate logistic regression was employed. (Pollock 2011: 212-213). The same model was run with three different dependent variables – restrictions, exclusions and judicial action.

Consistent with Buchman (2007) and the literature generally, judicial score was a key independent variable of interest. Similarly, in keeping with Buchman (2007) and the results of the preliminary examination of the behavior of magistrate judges, only Article III judges were included. Several additional variables were included as well. The first variable included was the interaction variable predicated on Buchman's (2007: 687) finding that individual plaintiffs fare worse when challenged by non-individual defendants. The second was the dummy for whether the case was published in the Federal Supplement, the official reporter for district level decisions, as tested in Buchman (2007) and suggested as potentially differing from unpublished decisions in Rowland and Carp (1996) as well as Taha (2004). The presence of different objections was added as a set of variables testing the impact of legal and cognitive impacts rather than political considerations. Whether the challenge was conducted at the summary judgment stage

was included consistent with Dixon and Gill (2001). Gender, minority and prior judicial service were included, in keeping with the general parameters of the behavioral model as set out in Boyd (2016). Similarly, in light of the potential for differences in approaches to these issues based on when the judge was appointed, the variable for whether a judge was appointed before *Kumho* expanded the scope of the gatekeeping function to all expert testimony was included. Based on the possibility that senior status may influence outcomes this was included as well. Consistent with the assertion in Dixon and Gill (2001) of potential circuit effects, as well as the results of initial investigation into the impact of location, dummy variables for circuit of origin were included. The reported results use the Seventh Circuit as the reference category, with dummy variables for cases originating in the Ninth and Fourth Circuits included in the model.

Table 27 reports the results of the model using restrictions on testimony as the dependent variable:

	B	S.E.	Wald	Sig.	Exp(B)
Judge Score	.486	.313	2.414	.120	1.626
Corp. Chall. of Ind. Pl.	-.031	.284	.012	.914	.970
F. Supp.	1.216	.422	8.321	.004**	3.374
Qualifications Chall.	.021	.285	.005	.941	1.021
Method Chall.	-.907	.299	9.197	.002**	.404
Relevance Chall.	.761	.285	6.984	.008**	2.141
Scope Chall.	1.967	.286	47.479	.000***	7.153
Summary Judgement	-.461	.316	2.127	.145	.630
Judge Gender	-.161	.373	.186	.666	.852
Judge Minority	.050	.486	.010	.919	1.051
Prior Judicial Service	-.447	.312	2.058	.151	.639
Appointed pre- <i>Kumho</i>	-.351	.560	.393	.531	.704
Senior Status	-.119	.594	.040	.842	.888
4th Circuit	-.058	.470	.015	.901	.943

9th Circuit	.760	.319	5.662	.017*	2.139
Constant	-1.103	.480	5.275	.022	.332
N	351				
Chi-Square/ Significance (a)	109.123/ .000				
Change in -2LL	349.160				
Cox and Snell R-Square	.267				
Nagelkerke R-Square	.367				

(a) Degrees of freedom = 15

*Significant at the .05 level

**Significant at the .01 level

***Significant at the .001 level

Table 27: Effects on Likelihood of Restriction; Logistic Regression Coefficients and Model Summary

Table 27 supports the conclusion that several different factors influence the decision of whether to limit testimony in some way. Three of the four categories of potential challenges to testimony bore a statistically significant relationship to the question of whether a judge would restrict testimony. The Ninth Circuit showed a significant relationship at the .05 level when examined in reference to the Seventh Circuit.¹² Publication in the Federal Supplement also bore a significant relationship to action at the .01 threshold.

The results relating to the various categories of challenge to expert testimony are especially revealing. They support the operation of heuristics consistent with the tenth hypothesis. Specifically, judges seem to be much more willing to uphold challenges that relate to issues they address in every case and tend to avoid or reject more esoteric questions such as understanding expertise and especially assessing methods. This in turn

¹² In an alternate iteration with the Fourth Circuit as the reference the variable for Ninth Circuit cases was slightly above the .05 level required for statistical significance.

is also consistent with the findings of Gatowski, et al (2001: 443-448) that although judges attempt to embrace *Daubert/Kumho* they often lack the knowledge necessary to address methodological issues.

Both the sign and significance relating to methods challenges lent further support to the tenth hypothesis regarding the potential impact of heuristics. Judges, being unfamiliar with methods and urged by the Supreme Court in *Daubert/Kumho* to default in favor of admission, appear to be reluctant to take action based on alleged problems with methodology in all but extreme cases. Despite methods challenges being the most frequently brought challenge in the sample, the existence of such a challenge was associated with a substantial reduction in the likelihood of a judge restricting an expert's testimony. Specifically, when all other variables are held at their sample mean¹³ the presence of such a challenge reduces the probability of a restriction by 21 points, from 48 to 27 percent.

Conversely, when judges are presented with the type of question that is more familiar, they appear to be more willing to sustain a challenge and restrict the expert. Both relevance and challenges to the propriety of an expert's testimony showed statistically significant and robust effects on the likelihood of a restriction being granted. Challenges to the scope of testimony had especially robust effects. While challenges to relevance enhanced the probability of restriction 18 points, from 26 to 44 percent,

¹³ In this context, the mean value for judge scores was calculated from the entire range of DW-Nominate scores reflected in the sample and their relative frequency among issued rulings. The sample means for all other variables was calculated as the proportion of each dummy variable coded positively in the sample in keeping with Pollock III (2011: 230-234).

challenges to the scope of a witnesses' testimony raised the probability of a restriction a tremendous 44 points, from 20 percent to 64 percent.

Again, this finding is consistent with both the qualitative analysis and prior literature on expertise and judgment. Judges address questions of relevance on a regular basis, as it is the fundamental question for the admission of all evidence. They similarly police the proper roles of witnesses, jurors, attorneys and themselves as a core element of triadic dispute resolution. In these settings judges are apt to have more confidence in their judgements and the appropriate standards on which to assess a violation. Attorneys are also more likely to be able to couch these objections in a manner that resonates with the judge's experience rather than being seen as technical or abstract arguments about issues that are at root foreign to their daily lives. Their relative expertise likely allows for rapid assessment and certainty in the correct outcome (Simon 1955; Kahneman 2011). This comfort might also permit a direct assessment of the merits of the objection rather than looking at peripheral factors such as industry standards and the rulings of other judges (Cooper and Neuhaus 2000).

The sign and direction, as well as the significance, associated with the variables for location support the second hypothesis about the uniqueness of the Ninth Circuit. This in turn provides support for the value of law in moderating judicial conduct. Both the Ninth and Seventh Circuits are split relatively evenly between judges who are appointed by Democratic and Republican Presidents. Judges in both articulate the *Daubert/Kumho* standards in similar ways. Yet the presence of a case in the Ninth Circuit, when the Seventh Circuit is used as a reference case, increases the probability of

a restriction by 17 points, from 25 to 42 percent. Moreover, although it does not quite achieve statistical significance the Ninth and Fourth Circuits have a very similar relationship. Whether this would extend to all judges in the Ninth Circuit, or if a similar split would be shown with other circuits, is a question worth developing further.

Finally, the significance of inclusion in the Federal Supplement supports prior literature on what leads judges to publish their work. Publication in this reporter is primarily at the discretion of judges (Taha 2004: 4). Much like publication in any other field, it tends to be skewed in terms of conduct that appears noteworthy – indeed, a lack of “interest to the legal community” is one of the few grounds for not publishing an opinion submitted by a judge for publication (Taha 2004: 4 n.4). Granting a motion to restrict testimony likely stands out as opposed to finding that there is nothing wrong with an expert’s proposed testimony that followed industry norms and standards. At the same time, unlike the other factors examined here, inclusion in the Federal Supplement likely cannot be seen as causal of outcome as the decision to publish is typically undertaken after a judge is well aware of the outcome of the challenge.

	B	S.E.	Wald	Sig.	Exp(B)
Judge Score	.777	.310	6.293	.012*	2.176
Corp. Chall. of Ind. Pl.	.195	.276	.498	.480	1.215
F. Supp.	.207	.386	.287	.592	1.230
Qualifications Chall.	.124	.281	.194	.660	1.132
Method Chall.	-.243	.306	.627	.428	.785
Relevance Chall.	.361	.272	1.760	.185	1.435
Scope Chall.	-.467	.306	2.336	.126	.627
Summary Judgement	.330	.288	1.314	.252	1.391
Judge Gender	.463	.353	.1722	.189	1.589
Judge Minority	-.137	.457	.089	.765	.872
Prior Judicial Service	.443	.306	2.092	.148	1.557

Appointed pre- <i>Kumho</i>	.357	.495	.519	.471	1.429
Senior Status	-.195	.536	.133	.716	.823
4th Circuit	.135	.450	.091	.763	1.145
9th Circuit	.949	.333	8.117	.004**	2.583
Constant	-2.620	.516	25.787	.000***	.073
N	429				
Chi-Square/ Significance (a)	23.941/ .066				
Change in -2LL	382.870				
Cox and Snell R- Square	.054				
Nagelkerke R-Square	.089				

(a) Degrees of freedom = 15

*=significant at the .05 level

**=significant at the .01 level

***=significant at the .001 level

Table 28: Effects on Likelihood of Exclusion; Logistic Regression Coefficients and Model Summary

As shown in Table 28, when the question turns to exclusion versus restriction, far fewer variables in the model demonstrate statistical significance and the model as a whole does not suggest it represents a meaningful improvement over random chance in light of the lack of statistical significance of the chi-square value and the low values for pseudo r-square. (Pollack III 2011: 228-230). With these caveats, two factors - partisanship and circuit of origin – both achieve statistical significance and appear to have a similar degree of impact. Although this suggests Republican judges and judges in the Ninth Circuit¹⁴ are both marginally more likely to exclude an expert witness, these differences pale in comparison to the overwhelming tendency in favor of admission regardless of other location or individual judicial factors. This in turn suggests that any

¹⁴ Notably these findings remained the same when the Fourth Circuit was used as a reference case for location instead of the Seventh Circuit.

claims of the impact of partisanship on district court judges should be treated cautiously, especially when (as here) the model has limited explanatory power.

The fact that partisanship is one of only two categories of variable to demonstrate a significant relationship with exclusions is intriguing on its face. The impact of moving from a mean score for Democratic appointed judges to a mean score of Republican appointed judge is an increase in the likelihood of an exclusion of approximately 9 points. This is approximately two-thirds of the change seen by Buchman (2007: 686) but is consistent in terms of direction. By itself, this would appear to provide support for the fourth hypothesis, drawn from work on judicial attitudes, that judges appointed by Republican judges are more willing to exclude experts. This in turn has a number of possible justifications, including but not limited to the Republican concern with tort reform and skepticism of the value of certain types of expert witnesses. (Kagan 2001 157-160; Schwartz and Silverman 2006: 226-231)

This relatively small shift in probabilities and the overall low probability of exclusion together indicate that the first hypothesis – that judges follow the legal model by defaulting to admission of expert witnesses - is supported by the evidence. The probability of exclusion when set at the mean judicial common space score of Democratic nominees and all other variables are held constant at the sample mean is 12 percent; for the mean judicial common space score of Republican nominees it is 21 percent. Consistent with the qualitative data and frequencies, it appears to be safe to say that judges are acting consistently with both *Daubert/Kumho* and the precepts of Rule 702 in a broad sense and reserving exclusion for exceptional cases.

This finding stands in stark contrast with Buchman's (2007: 688) assessment that *Daubert/Kumho* has not impacted judicial behavior. He observed much higher exclusion rates than were present in this study. Those rates were also inconsistent with Dixon and Gill's (2001: 271-272) finding that approximately one quarter of evidence was found unreliable before and after *Daubert/Kumho*. As noted above, Buchman's case selection method of looking for cases that used the term "tort" within the text of the decision may well have skewed his data set towards a more contentious set of disputes such as toxic torts or mass torts.¹⁵ That in turn would help explain his extraordinarily high rejection rate of expert testimony proffered by individual plaintiffs and challenged by corporations— a pattern which was not reflected in this dataset.

Additionally, this small shift in probability suggests a high degree of agreement amongst all judges on which cases do and do not warrant exclusion. Indeed, this degree of variance would likely be seen as the natural tendency of human nature in other contexts, with no assumption that it is the result of policy preferences or conscious choice (Galanter 1993: 90-91; Kahan, et al. 2016: 391-397). Overemphasizing the significance of a political difference of this size seems particularly problematic given the small improvement over chance this model represents based on the pseudo r-square scores

¹⁵ It is also noteworthy that Buchman identified less than 800 rulings on an individual expert basis in over 20 years on what was supposed to be a comprehensive national search for tort cases, yet the present data set covering only 9 out of 94 Federal district courts for six years identified over 400 rulings by Article III judges and over 500 total, the majority of which would be classified as tort disputes. Some portion of the disparity is attributable to the rise in filings over the intervening years, but the disparity suggests Buchman's dataset was not completely representative of how judges address expert witness testimony in light of *Daubert/Kumho* even in tort cases.

(.054 and .089) it returned. Thus, while partisanship is statistically significant, and the fourth hypothesis is supported, it is a distinction that is far from outcome determinative.

In fact, it appears that the location of a case says as much about the likelihood of exclusion as whether it is heard by a Democratic or Republican judge. When judges in the Ninth Circuit are compared to a reference sample from the Seventh, as a body they appear far more likely to exclude an expert. Here, when all other factors are held at their means, the shift to the Ninth Circuit as the circuit of origin increases the probability of exclusion from 11 percent to 24 percent. This supports the second hypothesis, which also held true in the context of restrictions. Likely owing to the differentiated case law of the Ninth Circuit, there is an appreciably higher likelihood a judge will restrict or exclude an expert's testimony. This in turn suggests that studies on judicial decision making at the trial court level should be mindful for the potential for the sort of circuit level distinctions noted in Songer, Sheehan and Haire (2003: 125-126) rather than assuming that judges throughout the country can be treated as an undifferentiated cohort as done in most studies.

	B	S.E.	Wald	Sig.	Exp(B)
Judge Score	.785	.265	8.737	.003**	2.192
Corp. Chall. of Ind. Pl.	.047	.230	.041	.839	1.048
F. Supp.	.945	.351	7.244	.007**	2.572
Qualifications Chall.	.013	.233	.003	.956	1.013
Method Chall.	-.727	.261	7.764	.005**	.483
Relevance Chall.	.670	.232	8.364	.004**	1.954
Scope Chall.	1.319	.246	28.735	.000***	3.739
Summary Judgement	-.028	.249	.012	.912	.973
Judge Gender	.142	.298	.225	.635	1.152
Judge Minority	-.065	.388	.028	.867	.937
Prior Judicial Service	-.107	.250	.183	.669	.898

Appointed pre- <i>Kumho</i>	-.406	.437	.862	.353	.666
Senior Status	.013	.476	.001	.979	1.013
4th Circuit	-.094	.371	.064	.800	.910
9th Circuit	.812	.258	9.928	.002**	.973
Constant	-.760	.400	3.610	.057	.467
N	429				
Chi-Square/ Significance (a)	93.311/ .000				
Change in -2LL	500.381				
Cox and Snell R-Square	.195				
Nagelkerke R-Square	.261				

(a) Degrees of freedom = 15

*=significant at the .05 level

**=significant at the .01 level

***=significant at the .001 level

Table 29: Effects on Likelihood of Judicial Action; Logistic Regression Coefficients and Model Summary

One of the concerns raised by the results shown in Tables 27 and 28 in isolation is that either could lead to a distorted image of judicial action as a whole. If judicial action is defined solely in terms of exclusions, it can lead to a conclusion that judges do not apply *Daubert/Kumho* with much rigor. That was certainly the conclusion of Helland and Klick (2011), when they compared the sheer number of experts at trial in state and Federal jurisdictions. No matter what certain advocates urged, by its own terms *Daubert/Kumho* was intended to liberalize rather than constrain expert testimony. The obligation imposed on judges is to be mindful of the potential for overreach, something that they apparently believe can be accomplished through restrictions in most instances rather than exclusions. If only restrictions were considered, though, it would be just as distortive as it would completely ignore the impact of partisanship and suggest that judicial attitudes play no role in judicial conduct.

To address this potential oversight a variable encompassing both restrictions and exclusions was coded in this data set and used as the dependent variable in the same model. It trades granular knowledge of either particular type of action for a better overview of the entire picture of judicial conduct. As shown in Table 29, it manages to capture the full range of significant variables including partisanship, categories of challenge brought, location¹⁶ and publication. There is some loss in predictive power relative to the model for restrictions but a significant gain over its use in the context purely of exclusions. This would appear to be an improvement in truly understanding influences on judicial decision making in the context of this and similar contexts where rulings that fall between an absolute grant and absolute rejection of a challenge are possible.

Variable		Change in Probability	
	Restriction	Exclusion	Judicial Action
Judge Score Full Effect D. to R.	---	12	23
Judge Score – Median D. to R.	---	9	18
Method Chall.	-21	---	-18
Relevance Chall.	18	---	17
Scope Chall.	44	---	32
9th Circuit	17	13	20
F. Supp.	29	---	23

Difference in probability determined by comparing mean at values of 0 and 1 for all dummy variables. For full effect of partisanship probabilities calculated at maximal Democrat (-.45) and Republican (.73) judicial common space score; for median difference, probabilities calculated at median Democrat (-.379) and Republican (.598) judicial common space score.

Table 30: Probability Differences for Variables Showing Statistically Significant Effects

¹⁶ In this model the variable for the Ninth Circuit is statistically significant regardless of which circuit is used as a reference, but the variables for the Fourth and Seventh Circuits were only statistically significant when compared to the Ninth.

Table 30 illustrates the value of this perspective. A direct comparison of the relative impact on the probability of each statistically significant variable in this model shows that most of them have similar absolute magnitudes between 17 and 23 percent. The one exception is the impact of challenges that go to the proper scope of a witnesses' testimony. Judges appear to be particularly sensitive to the potential for experts and their counsel to overstep their bounds, which is both in keeping with legal norms given the adversarial nature of the system and the particular concerns about the potential for experts to stray from appropriate testimony set forth in *Daubert/Kumho* and other cases. At the same time, this study shows that judges primarily react to such violations through restrictions designed to keep experts within acceptable boundaries, while not transgressing the abiding respect most judges profess for the value and capacity of the adversarial system.

As summarized in Table 31, ultimately the quantitative analysis largely supported the key themes identified by the qualitative analysis. Judges do appear to default in favor of admission and seek to avoid exclusion except in extreme cases. Judicial decisions show that in the more common situation of challenges that result in exceptions rather than exclusions there appear to be some efforts at cognitive simplification at work. Specifically, they appear much more willing to step in and police the appropriate bounds between attorneys, experts, judges and jurors than they are to engage and analyze methods, in keeping with their own relative degree of training and familiarity with the two subjects. Thus, both the legal model and cognitive influences do seem to be influencing judicial conduct.

Hypothesis	Data Supp.(a)	Reject Null Hypoth.(b)
H1. Consistent with the precepts of the legal model, judges as a class will show a common tendency to admit all experts.	Yes	Yes
H2. Judges within the Ninth Circuit will uphold challenges to expert testimony at a higher rate than their colleagues in other circuits.	Yes	Yes
H3. Judges within the Fourth Circuit will be less likely than their colleagues to uphold objections based on the qualifications of an expert or the relevance of their testimony.	No	No
H4. Judges appointed by Republican Presidents will be more likely to uphold challenges to expert testimony than those appointed by Democratic Presidents.	Yes	Yes
H5. Judges appointed by Republican Presidents will be more likely to uphold challenges to expert testimony made by corporate or public defendants sued by individuals than those appointed by Democratic Presidents.	No	No
H6. Female judges will be more likely to uphold challenges to expert testimony.	No	No
H7. Judges on senior status will be less partisan than active judges.	No	No
H8. Judges will be more likely to exclude expert testimony at the summary judgment stage.	Yes	No
H9. Judges will be more likely to restrict rather than reject the testimony of an expert witness even when a legal challenge is upheld.	Yes	Yes
H10. Consistent with the concept of heuristics, judges will be more likely to uphold objections that are similar to rulings they are regularly called on to make.	Yes	Yes
H11. Judges appointed before the completion of the <i>Daubert/Kumho trilogy</i> will be less likely to uphold challenges to expert testimony.	Yes	No

(a)Data support defined as consistent findings in keeping with hypothesized relationship

(b)Reject null hypothesis defined as one or more findings consistent with hypothesized relationship and statistically significant at the .05 level

Table 31: Summary of Support of Hypotheses

At the same time the quantitative analysis added a dimension the qualitative could not supply. It demonstrated that judges appointed by Republican Presidents have a greater willingness to exclude expert witness testimony than their counterparts appointed by Democratic Presidents. Although this is not so pronounced as to be outcome

determinative in most cases, it does suggest that even in areas of law that are not evidently politically salient this distinction can still matter. It is likely more pronounced in subpopulations of *Daubert/Kumho* challenges in more politically salient contexts, but that is an investigation that is beyond the scope of the present project.

The analysis also provided evidence for the value of institutional considerations when assessing district court judges. By controlling for location, it was discovered that although Democrat and Republican appointees in the Ninth Circuit follow the split predicted by attitudinal models, it is less pronounced than in other locations. Indeed, local impacts mediate the effect to the point that Democrat appointees in the Ninth Circuit would act more like their Republican counterparts than they do Democrats in the Seventh Circuit. This suggests much more study should be done into if and when circuit effects mediate the impact of partisanship.

In fact, one of the key findings from the quantitative study is that a number of potentially relevant factors warrant further study. Although it does seem to rather soundly reject the idea that judges in the Fourth Circuit are hostile to challenges based on the qualifications of the witness or relevance of testimony despite their invocations of *Kopf*, many other relationships showed consistent results but did not reach statistical significance. The fact that judges consistently excluded experts at a higher rate in the context of summary judgment motions and yet granted fewer restrictions speaks both to strategic considerations and cognitive burden minimization and may well prove to be statistically significant in an expanded dataset. Similarly the tendency of judges who were appointed before the *Daubert/Kumho* trilogy was complete to have a lower rate of

action overall, but not exclusions, raises some interesting questions about the process of how judges acquire and maintain their views of the appropriate scope of their role.

In sum, the quantitative study thus suggests a broad array of influences can have an impact on judicial decision making. At the same time, these influences work on the margins of outcomes in the context of *Daubert/Kumho* and against the backdrop of what appears to be a near consensus on how to resolve disputes over expert testimony. Both this background consensus and the other identified influences raise some significant questions about how legal policy and precedent should be crafted to ensure alignment of judicial conduct with the intent of any legal standard. They also speak to some fundamental limits on how far policy makers can consciously alter the practices of judges without first shifting norms within the broader legal community and the broader society. Trying to meaningfully place these findings as well as the themes identified in the qualitative analysis in context and suggest some practical applications of these insights is the task of the next and final chapter.

Ch. 5 - Conclusion – Seeing Judicial Decision Making as a Mosaic and Plotting a Path Forward

As shown in both qualitative and quantitative analyses, a range of factors seem to influence judicial decision making. At one level, the outcomes are consistent with the central goal of *Daubert/Kumho* with regards to expert testimony – most experts are admitted, although many have limitations placed on their testimony. At the same time, the path that judges use to get to admission appears to rely less on rigorous assessments of scientific or technical methods and much more on the norms of the adversarial process as well as socially constructed standards regarding various subjects and witnesses, raising questions about whether they would be so clearly in keeping with the law if its requirements were less forgiving. Additionally, both location and partisanship seem to impact this tendency, albeit neither to a degree that would change the conclusion that judges default to admission. Even when a challenge is successfully lodged against an expert, restriction rather than exclusion is favored, further underscoring the conclusion judges interpret and apply the law to favor admission. How judges structure their analyses of challenges, their reliance on industry and legal norms as substitutes for independent examination, and the relative success rates of different types of challenges all indicate heuristics are being employed to ease the cognitive burdens associated with this task. Additionally, there is suggestive evidence that strategic considerations come into play based on when challenge arises in the course of a case and the influence of intergenerational changes in role perceptions, but neither of these influences has been demonstrated to meet the standards for rejecting the null hypothesis.

Thus, even the seemingly mundane task of pre-trial rulings reveals itself to be a complicated process. While the command of the law to admit experts in most cases appear dominant, many others are also at work – and still others might be shown to have an impact, or the magnitude of impact might shift, if the subject matter were restricted to particular types of cases. Similarly, expanding this study to other courts or a different time frame might reveal a very different pattern of outcomes. This complexity suggests that advancing the field of judicial decision making significantly will require scholars to acknowledge and account for this nuance while retaining perspective on how individual studies relate to the whole of this scholarship as well as the broader societal context in which judges make decisions. Otherwise the field is unlikely to transcend the current divisions amongst legal scholars and the various social scientists that adhere to the myriad variations of the judicial attitude, behavioral, strategic, institutional and cognitive models of judicial behavior.

Beyond suggesting a path forward for following up on this study and judicial decision making scholarship generally, the results of this study also point to the complexity of designing meaningful policy changes within the legal field.

Daubert/Kumho was intended to liberalize the admission of expert testimony, but also assure that experts and their methods received meaningful scrutiny. The former has been accomplished, but the expansive discretion given to judges to decide how to assess the latter and the lenient standard of review used by most circuits may be thwarting the latter. The lasting impact of *Daubert/Kumho* was also likely blunted by the way the Supreme Court sought to conform the gatekeeping function to pre-existing norms about the scope

and nature of the adversarial process. While these efforts likely made the cases more readily accepted by judges, they appear to have reinforced a belief in the efficacy of cross-examination and contrasting accounts as a way to clarify disputes about expert testimony that is contrary to empirical evidence.

All of this suggests that any policy maker in this area, whether they are a judge or legislator, should carefully consider how proposed changes are likely to fit or conflict with existing norms and expectations and craft the policy accordingly. Any significant change that conflicts with existing norms may require policy makers to restrict the discretion of the parties tasked with carrying out the new policy to limit shirking or overreliance on existing norms. Creating lasting change may also require policy makers to go beyond simple, vertically imposed rules to a more entrepreneurial stance, potentially including public explanation and efforts at incorporating norm changes in the socialization of law students into the bar. This last task may also require broader efforts at shaping societal expectations of the legal system to minimize conflict between underlying societal norms and those embraced by members of the legal community.

Finally, this study suggests that practitioners should adopt certain approaches to maximize the chances of successfully challenging an expert witness. If exclusion is the goal, it is not enough to point to the fact the expert lacks ideal credentials, failed to adhere strictly to industry standards or appears to go beyond what is strictly necessary to a party's legal case. Such challenges are apt to be met by the court with an invitation to explore it on cross-examination at trial, or a retort that the issue "goes to the weight, not the admissibility" of the evidence. Rather, the proposed expert and their testimony must

be portrayed as being significantly deficient or deviating severely from expectations in the field that the testimony relates to, ideally by the expert's own admission. At the same time, if the objective is simply to rein in testimony, the best route appears to be to explain either its irrelevance or suggest a way in which it transgresses the appropriate role of an expert. Regardless of the goal however, practitioners would be advised to consider if the effort of the challenge is justified given the tendency of judges to admit expert testimony and their underlying belief in the value of cross-examination and contrasting arguments to address the vast majority of issues related to expert testimony.

A. There are Several Extensions and Refinements of this Study that Deserve Attention

Although the findings in this study appear to be backed by robust qualitative and quantitative evidence, there are still reasons to view its conclusions with caution. In part this is tied to the purposive sampling case selection process used in this study. Whether the positive findings can be replicated when applied more generally remains to be seen. Similarly, it may be that the failure to find meaningful distinctions linked to gender, race and ethnicity or the time of appointment to the bench may be a function of relatively small numbers of subjects meeting each of these criteria. The findings on the effect of the timing of a challenge to an expert witness were consistent but fell short of statistical significance, suggesting this is another area that could be meaningfully explored in greater depth. Moreover, the review in this case has identified several factors that were not directly linked to the focus of this study and were not capable of meaningful analysis with the present dataset. Each of these issues in turn support further explorations of both

district court judicial decision making generally and the field of expert witness challenges particularly.

This study was consciously designed as exploratory rather than definitive in nature. For example, purposive sampling was employed to ensure that certain variables of interest like circuit level distinctions were preserved (Seawright and Gerring 2008: 300-301). While this choice was justifiable in light of identified unique features of Ninth Circuit precedent, it necessarily inhibits the confidence that can be ascribed to these results. The overwhelming tendency to admit expert witnesses seems likely to hold in all cases, but other findings may be artifacts of the research design. In particular, the tendency of Democrat appointed judges in Oregon and Washington to be more critical of experts than any other cohort outside of this region is consistent with the impact of Ninth Circuit precedent but not definitive. It may be due to unique features of these particular courts, or a similar idiosyncratic pattern in the courts from the Seventh and Fourth Circuits that were used as references. Expanding the study to embrace the entirety of these circuits would likely provide greater confidence but would require careful efforts at randomization and equalization of sample size given the Ninth Circuit's much greater geographic scope, number of judicial personnel and collective caseload. It would also be beneficial to add further comparisons such as to the Fifth or Second Circuits, although doing so without making the process of data evaluation cumbersome might require limiting the analytical focus significantly.

The findings of no statistically significant correlations between case outcomes and behavioral distinctions linked to gender, race or ethnicity and relative time of

appointment also should be taken with caution. Ultimately there simply were too few judges who met these markers to have much confidence in these findings. Racial and ethnic diversity among Article III judges in particular was so limited in this sample that any result must be met with skepticism. One possible way to ameliorate this issue would be conscious oversampling or restriction to such candidates, but it would likely sacrifice the ability to control for circuit effects. A matched pairs sample of particular male and female judges or white and minority judges might also be a possible means of exploring these issues in particular, although again this raises questions about how to retain the ability to control for several of the other significant influences identified in this study such as partisanship given the correlation of Democratic appointments and female and minority judges.

The potential for strategic action by judges also appears to merit further exploration. The data in this study showed that regardless of the controls used the rate of exclusions of experts went up when they were brought in the context of a summary judgment motion relative to other phases of a case and the rate of restrictions went down. The failure of this finding to achieve statistical significance likely is a matter of the relative infrequency of such situations rather than an indication the relationship is due to random chance. Thus, expanding the dataset would likely lead to a more definitive answer on the nature and scope of strategic considerations by district court judges in making this sort of determination. Additionally, in a study focused specifically on this question it would be beneficial to see if the rise in exclusions tracked a rise in resolutions

short of trial, as that is the likeliest incentive for judicial exclusion of expert witnesses at the summary judgment stage.

In addition to these refinements on the existing study, several other interesting observations in this data warrant further investigation although the parameters of these investigations are unclear at the moment. First, there was an overwhelming tendency for judges to reject challenges by saying they properly went to the weight of evidence or were reserved as subjects of cross-examination. Although this supported the qualitative finding that judges believe in the adversarial system and are reluctant to intervene outside of extreme cases, they were not really capable of quantitative assessment as influences on judicial decision making as structured in this study as they were universally associated with rejection of challenges. A more fruitful investigation might explore how prevalent such statements are in other areas of evidentiary rulings to in turn see if the tendency towards limited intervention is more universal or uniquely linked to the subject of expert witnesses.

A second suggestive thread in the data that defied direct examination in this study was the prior work experience of judges. Prior judicial service was suggested as a potential basis for differentiated handling of expert witness challenges but it did not hold up in even initial testing. However, in compiling this data about judges it emerged that minority and female candidates have a more constrained pathway to the Federal judiciary than white males. The latter are equally likely to have served in government as attorneys, been previously elected or appointed as judges or been in private practice before appointment to the Federal bench. Female and minority candidates appeared to be

predominantly selected from the first two career paths, although the distinction was small and potentially illusory. A wider study of this tendency, controlling for the fact that the proportion of women and minority judges has grown substantially in the last two decades, might be informative both of questions of systemic bias in the selection system and whether previously identified gender associated traits in decision making are related to this constrained path to selection.

A third area identified in this study for future exploration is the development of judicial boilerplate. Most judges appeared to have a form legal analysis section for addressing challenges to expert witnesses, but they are not unique or static. Judge Herndon's rulings in the *In re Yasmin* litigation added and subtracted paragraphs based on the field of expertise considered. Judge Flanagan's rulings typically, but not universally, referenced *Kopf* which appeared to signal different assessments by her of certain subjects. Moreover, reading judges who sat in the same district often showed some measure of mutual exchange and adoption of standards, although these were rarely explicitly acknowledged. Studying the development of these sections in more depth might supply insights into how judges minimize cognitive burdens as well as how they work as a community despite technically sitting on all matters individually. However, doing so would likely require district judges to openly discuss practices that they are often reluctant to comment on.

As with most studies, this one raises as many or more questions than it answers. There are issues of generalizability that require further exploration to answer. There are suggestive but inconclusive results that merit additional scrutiny. There are tantalizing

bits of information that could not be tested with the current data but with some development could lead to significant insights. District court judicial decision making, both generally and specifically regarding expert witness testimony, is a field replete with possible further opportunities for study. How and if the results of those studies will help create an improved and meaningful understanding of this field, though, depends on whether scholars can adopt a more holistic perspective or will remain attached to the idea that one model or other is a superior explanation of judicial conduct.

B. This Study Shows Judicial Decision Making is a Mosaic – and Scholars Should Embrace its Complexity and Nuance

While this study is itself only a beginning of a potentially lengthy investigation of judicial decision making, it demonstrates the multi-faceted nature of the process. Unlike most studies in the field, it revealed not only a broad range of influences at work but also how they were largely compatible with each other. This suggests that the tendency of scholarship in this field to focus on particular explanatory models is potentially creating a distorted image. Like a mosaic, a fuller picture emerges only with perspective and appreciation for how components complement each other.

As Epstein (2016) and Rachlinski and Wistrich (2017) both document, a broad range of approaches have been taken to studying judicial decision making. Everything from political identification to gender to speech patterns and the investment portfolios of judges have been examined with mixed outcomes. Many studies can find one or more variables that have a statistically significant relationship with the outcome of legal decisions. This can lead to an impression that essentially everything about a judge and a case might matter.

Yet at another level the collective body of this scholarship frequently seems to be less than the sum of its parts. In part this stems from the fundamental lack of epistemological agreement on what counts as a meaningful result, rather than just a statistically significant one. Similar rates of variation are trumpeted as profound in one context and as essentially meaningless in another. As a field, judicial decision making has yet to consolidate into a period of “normal science,” consistent with Kuhn’s (1996: 15, 160-165) view that social science is generally too riven with options and competing perspectives for such moments.

Another key barrier to evaluating the contributions of judicial making scholarship is the tendency of most studies to focus as narrowly as possible in terms of the explanatory lens brought to bear on subjects. Very few studies seek to explore more than one possible explanation for judicial conduct. While this sharpens the focus of a particular article it has a splintering effect on the literature as a whole. A number of competing claims can arise based on similar evidence as a result of minor changes to parameters, with the conclusion of what matters predicated on which elements the investigator chose to emphasize rather than a broader perspective. While that tight focus may be good for achieving publishable results, it leads to a loss of perspective, as noted by Epstein and Knight (2013: 12-13).

Even the most commonly investigated relationships are subject to fundamentally contradictory interpretations of the same results. Focusing just on the current study, a supporter of the legal model could justifiably say that the dominant tendency of judges to admit experts dwarfs the magnitude of differences between judges based on their partisan

alliances or any other factors. They could look to studies discussed by Galanter (1995: 90-91) and Kahan, et al.'s (2016: 391-397) experimental results and say the difference between judges appointed by Republicans and Democrats is nothing more than a representation of the natural variation in any field of assessment. Indeed, the gap is smaller and more consistent than many of those in other fields, suggesting a higher degree of isomorphism and validation that the rule of law dominates judicial conduct. Moreover, they could compare the relative rate of exclusions by Democratic appointed judges in the Ninth Circuit with any group of judge from the Seventh or Fourth Circuits and argue that this shows the influence of partisanship is illusory and really what matters is local circuit precedent.

A fundamentally different conclusion might be arrived at by an adherent to the perspective that all judges are motivated policy actors. The fact that there is a persistent gap between judges based on their partisan identification, even at the lowest rung of the Federal judiciary and on seemingly mundane matters could be hailed as evidence that politics are inescapable in all legal decisions. Indeed, this is the conclusion Buchman (2007: 687) chose to highlight although he admitted it was not the largest influence he found. Similarly, although not statistically significant the fact that this partisan divide persists even after a judge takes senior status would likely be seen as a target for further exploration, as it potentially speaks to just how deep-seated the divide is. That relatively higher rate of exclusion by Democrat appointed judges in the Ninth Circuit, moreover, would not pose a challenge to this conclusion, as there is also a predictable gap fitting the expectations of the model between Democrats and Republicans within the Ninth Circuit.

A follower of the strategic/institutional model might see yet a third explanation. The nominal compliance with the overall goal of defaulting to admission while avoiding robust methodological analyses could be seen as a collective effort at strategic shirking of the more onerous elements of gatekeeping. The gap between Democrats in the Seventh and Ninth Circuits might speak not directly to law but to a strategic evaluation of just how far a Democrat appointee in the Ninth Circuit can be a sincere policy actor. The imposition in *Mukhtar* of a higher penalty for too casually dismissing a challenge to an expert may have been internalized as a check on their personal interests. Alternatively, the higher frequency of exclusions by all judges in the Ninth Circuit could be seen as a matter of these judges embracing the opportunity created by the Ninth Circuit's precedent on issues like causation evidence to maximize their personal goals of efficiently addressing cases through pre-trial rulings. Given the plethora of goals the strategic model has embraced, a myriad of plausible justifications could be offered (Epstein: 2016: 2041-2048).

For most of the last 40 years, moreover, the literature has been replete with studies whose outcome has arguably been overdetermined by case-selection or research design. Sometimes this is overtly acknowledged, as when Carp and Rowland (1983) consciously selected district court policy making on politically salient issues because they simply wanted to establish the existence of the phenomenon rather than its extent. Similarly, the selection of Supreme Court Fourth Amendment jurisprudence cases as the centerpiece of Segal and Spaeth's (2002: 320-330) study of judicial attitudes took an issue where the authors already knew there was a pronounced split on the Court and

examined its extent. Kahan, et al (2016: 353-354, 391-397) tout their experiment as overcoming case selection bias in Supreme Court studies, but their experimental design consciously nudged judges to follow the law and laypeople to study the facts, thereby ignoring the ambiguity and contestation inherent in many areas of law. As both Boyd (2016: 795-796) and Rachlinski and Wistrich (2017: 207-208) note, studies of judicial decision making are most likely to reveal tendencies linked to gender or race when cases are selected that are salient to women or minorities and not in other situations. These examples beg the question of just what these studies are trying to prove except to provide pre-determined ammunition for essentially contestable propositions about what matters more in judicial decision making.

As an alternative to these splintered and over deterministic efforts, it would seem to benefit scholars to try and take a more holistic and nuanced view of judicial decision making. There is a simple way to reconcile all of the findings in this study – by thinking of them as pieces of a mosaic or collage rather than competing at all. Like any other human being, a judge can and must satisfy multiple constraints simultaneously (Robbennolt, MacCoun, and Darley 2010). A judge can simultaneously be consciously committed to adhering to the law, conforming to local norms based on precedent or custom, taking the path of least cognitive burden so as to maximize their job performance and in so doing tending to take positions that are colored by the same traits that lead them to identify as a conservative or liberal. Bailey and Maltzman (2011) documented such a multitude of factors impacting the Supreme Court, and Epstein and Knight (2013) acknowledged the need for the field to take seriously the idea that there are motivations

beyond policy even in that rarified body. At the district court level, Rowland and Carp (1996) created a conceptual framework that embraced this view based on their own scholarship and that of authors like Cooper (1988) and Robbennolt, MacCoun, and Darley (2010) postulated a similar model drawn from psychology. Still, studies tend to frame the question of judicial decision making as one of competing rather than complementary factors, of discovering the key and previously unknown influence. It would appear to be well past time to accept that the humanity of judges necessarily means their decisions will be the result of a broad array of influences, often operating without conscious thought.

This is not to suggest scholars should never investigate individual influences. Identifying the areas in which different factors matter is an important contribution, as is mapping out the contexts in which those influences do and do not appear to have an appreciable impact. However, it would be beneficial for these studies to at least acknowledge that below the Supreme Court virtually every influence on judicial decision making linked to politics, behavioral traits, institutional constraints or cognitive influences is relatively small, context specific and unlikely to be outcome determinative in most cases (Epstein 2016; Rachlinski and Wistrich 2017). Indeed, almost every variation in the literature pales in comparison to the general consensus among legal actors, such as the 90 percent unanimity of appellate decisions (Kuersten and Songer 2014: 83) and the overwhelming consistency of outcomes in the current study.

Moreover, the tone of much of this work tends to suggest that judges are not honoring the rule of law in light of these relatively small differences. This naturally tends

to be resented by judges, which in turn leads to a fairly dismissive attitude towards empirical studies of judicial decision making (Tamanaha 2009; Edwards and Livermore 2009). These responses will frequently note that the existence of different, equally legitimate approaches to legal interpretation within the adversarial system will almost always yield differentiated outcomes as to how best to resolve novel questions of law. Even the pattern of ideological breaks shown in many areas of Supreme Court practice can be reconciled with the rule of law through the fact that these are areas of “open texture” in which a judge’s views necessarily influence any outcome (Hart 1994:145). For judicial decision making to be anything more than a series of interesting observations lawyers and judges ignore, scholars would be advised to acknowledge when, as here, judicial conduct is consistent with the rule of law and to place the observed differences in perspective.

A similar measure of perspective would be beneficial in comparing the distinctions found among judges to those present in other fields or the general population. The split identified along ideological lines in this study, for example, compares favorably if imperfectly to intra-group variance in other fields of judgment (Galanter 1993: 90-91). It also is higher than the rate of consensus taken by Kahan et al (2016: 391-397) as evidence of a lack of political influence on judicial conduct. In studying the impact of heuristics on judges, Guthrie, Rachlinski and Wistrich (2007: 13-16) explicitly compared judicial performance on the Cognitive Reflection Test and similar tasks with other groups in the population. This type of comparison is typically missing but would be highly informative to understanding whether judges are deviating from the legal model or if in

fact legal training and socialization mediates the natural differentiation of the population at large. For example, studies have consistently shown that judges split on the issue of sexual harassment cases on gender and ideological lines (although the two tend to be collinear in light of the appointment practices of Democratic Presidents relative to Republicans). This is an area of known fissures in the general population along similar lines, but studies largely omit comparisons of the distinctions found among judges with those known variations. Providing that sort of comparative data point would create the basis for a clearer standard for when such variances among judges merit policy interventions to correct.

Ultimately, more than anything else this study demonstrates that judicial decision making at the district court level is subject to a broad range of influences. Although the outcome of this particular set of cases was driven primarily by legal expectations, institutional influences, partisanship and cognitive simplifications were all present as well. Focusing on any one influence would likely have led to a failure to identify the others, or to overemphasize the partial effects of partisanship while downplaying the fundamental influence of law. Advancing the field of judicial decision making will require scholars find ways to document their observations while endeavoring to place those findings in the broader context of the broad range of influences that can impact judicial decisions, as well as how judges are similar or different than the population at large.

C. Policy Makers Should be Mindful of Cognitive Influences in Crafting Legal Doctrine and the Need for Shifting Cultural Norms to Achieve Lasting Changes

In addition to its lessons for scholars studying judicial decision making, this study also points to the difficulties inherent in changing the behavior of judges through policy shifts such as rule amendments or new precedent. When these vertically imposed alterations create potential conflict with underlying norms of behavior, even a judge trying to follow the former is likely to find ways to reconcile it with the latter. This in turn suggests that policy makers, whether they are judges or legislators, should carefully consider how a new obligation or standard fits with norms. If there is a clear conflict, and compliance with new policy is desired, crafting it with limited discretion may be advised. Moreover, to make the change lasting, meaningful and in accordance with the intent of the policy maker will likely require a number of efforts at shifting expectations and norms over time.

The *Daubert/Kumho* line of cases created a new practice for judges. As Dixon and Gill (2002: 252) document, before this line of cases it was not clear what the standards for admission of expert testimony were between the “general acceptance” test of *Frye* and other circuit level interpretations. After *Daubert*, judges were tasked with “gatekeeping” to ensure that methods generally were reliable in a legal sense and that experts had applied the same degree of rigor in arriving at their testimony that they would if they were applying professional judgment and expertise outside of court. This was in addition to assessing if the expert appeared qualified and the testimony was relevant to the issues at hand.

Had the Supreme Court simply imposed this obligation, it is uncertain how cases would be decided. However, the Court did not couch this obligation in a vacuum. Both in *Daubert* and in *Kumho* the Court made clear that admission should be the rule and not the exception, with any sort close call being resolved in favor of admission and the workings of the adversarial system. This emphasis on liberalizing of evidence law was in turn picked up by the Advisory Committee in revising Rule 702 to codify the *Daubert/Kumho* line of cases. This “choice architecture” is likely the reason that the overwhelming majority of experts are admitted regardless of the location, gender, experience or partisan identification of the judge making the decision.

While the goal of a default to admissibility has thus been achieved, it is less clear that the focus of gatekeeping has remained where the Court apparently intended. This can also be traced to the Court’s own holdings and further interpretations of the cases by intermediate appellate bodies. The initial statement in *Daubert* that judges should undertake a rigorous evaluation of scientific methods and the provision of a Popperian set of criteria for evaluating them including falsifiability/testing, error rate, publication/peer review and general acceptance pointed to the need for judges to acquire a new skill set and truly understand methods in most cases. However, the Court promptly undercut its message by indicating the criteria were non-exhaustive and that no factor or factors was determinative. The holding in *Joiner* to give deferential review to both what factors district courts used to evaluate expert witnesses and their conclusions gave judges a further signal that they did not have to apply *Daubert* with a tight focus in most cases. The expansion of gatekeeping to all types of expert testimony in *Kumho*, accompanied as

it was with further statements that most issues with expert testimony could be addressed through the adversarial process and reassurances that the precise named *Daubert* factors did not have to be applied in all settings yet again suggested that district judges were not expected to really comprehend the details of expert testimony. Most circuit courts, outside of the Ninth Circuit's *Mukhtar* decision, also made the consequences of error relatively painless – if there was an insufficient basis for the ruling to admit or exclude, the remedy was in most cases to supply a more sufficient analysis rather than undo an intervening trial.

What judges were left with was an exhortation to be gatekeepers, but an incentive to keep the gate open consistent with their pre-existing norms of respecting the adversarial process as documented by Cecil and Willging (1993: 5, 20-21). Moreover, the flexibility of the actual examination provided judges with a means to avoid the cognitive burdens of understanding scientific and technical methods except in extreme cases. This, coupled with the natural human inclination to minimize cognitive burdens noted by Simon (1955: 101-105) and Kahneman (2011: 31-38), is the likely explanation for the finding in this study that judges typically disregard the *Daubert* factors as inapplicable. Instead of looking for analogous field specific, verifiable indicators of reliability judges frequently collapse the inquiry to a cursory examination comparing asserted methods to generic industry or legal standards. This may mean that a lot of suspect testimony is permitted simply because no one in a position to do so has the incentive to explore whether common practices in the field are yielding meaningful results. The risk of this is demonstrated by *Cascade Yarns* – there is no way to know

how many cases were wrongly decided because of faulty fiber analysis but given that the expert had created an entire business around his alleged expertise the number is likely non-negligible.

This risk points to the need for policy makers to carefully consider how proposed changes in legal standards or practices align with judicial training and engrained norms. Before *Daubert* was announced, many Federal district judges indicated that they did not appoint independent experts despite clear authorization to do so out of concern with interfering with the adversarial process (Cecil and Willging 1993: 20-21). The goals of *Daubert/Kumho* that judges are to ensure scientific rigor in a pre-trial proceeding ran up against that norm of party driven litigation. Once the Court made it clear judges could minimize or avoid that conflict altogether the potential for judges to adopt a process that really focused on the unfamiliar concept of scientific and technical methodology was greatly reduced. It was requesting a change in behavior that violated deep seated norms without backing failure up with meaningful sanctions or even acknowledging the conflict. The result – a system that follows the law as written but minimizes the deviation from pre-existing norms – was predictable.

Whether this result was preventable cannot be definitively known, but there are some examples that illustrate alternative outcomes. As Beebe (2006) documented, giving judges a multi-factored set of standards for trademark law gave each circuit and ultimately individual districts the freedom to pick and choose the elements that would be emphasized with a result that there is not one test but myriad ones. By comparison, the standards for granting a preliminary injunction have far less variability – the same four

factors are used in all courts. Not only is disregarding one grounds for reversal, but the *Winter* case makes clear that even lowering the threshold for satisfying the irreparable harm from a “likelihood” to a “possibility” is unacceptable. The Court could have required a similarly constrained set of standards if it truly wanted judges to examine methods closely.

Alternatively, the Court could alter the norms surrounding the adversarial process by acknowledging the documented problems with using adversarial processes to counter suspect expert testimony and trying to reform how cases are tried generally (McAuliff, Kovera and Nunez 2008; McAuliff and Duckworth 2010; Levett and Kovera 2008). This may sound outlandish, but the continuous change of legal norms is one of the hallmarks of the common law system. Periods of greater consensus on the Supreme Court have been followed by the current trend towards block voting on issues with clear conservative/liberal value splits and unanimity in less contentious areas (O’Brien 1996; Epstein and Knight 2013: 16-17). The relative value of different sources of legal authority have shifted over time, too - an emphasis on precedent before the 1930s was replaced by decades of near consensus deference to legislative action. (Bailey and Maltzman 2011: 87-100). Currently there is a more fragmented picture with the emergence of avowed textualists and originalists like Justices Scalia, Alito, Thomas and now Gorsuch, more textually driven decisions from Justices Kagan and Breyer as well, and a re-emergence of precedent as being valued to a greater degree than Congressional intent.

These tendencies have filtered down to the appellate level. Regional splits used to dominate decision making. Now circuit level distinctions of the type found in this study have largely supplanted those patterns (Songer, Sheehan and Haire 2003: 125-126). District courts have gone from a primary focus on trial oversight to docket management and settlement facilitation (Hornby [2007]2013: 113-114; Boyd 2013: 194). There is no inherent reason norms surrounding the treatment of expert witnesses could not change as well.

At the same time, changing norms is typically more gradual and generational and requires a measure of policy entrepreneurship judges frequently disclaim. Examples of these processes abound in the law. The concept of a right of privacy was initially articulated in the Supreme Court in a dissent by Justice Brandeis in *Olmstead v. United States* (1928) 277 U.S. 438, building on his prior academic work. A lonely voice at the time, over the next several decades the concept if not his particular version of it became one of the defining principles for understanding everything from the validity of warrantless searches to the right of women to seek abortions. The growth was gradual, but at this point it would be hard to convince the public or all but the most extreme textualist judge that the right to privacy is not at the heart of our Constitutional liberties.

Similarly the history of strict liability for product defects shows how even one judge, acting as a policy entrepreneur, can lead a Kuhnian revolution in legal norms. Justice Roger Traynor of the California Supreme Court first set out the conceptual framework of strict liability in product defect cases in a concurrence in *Escola v. Coca-Cola Bottling Co* (Cal. 1944), 150 P. 2d 436. Traynor's reward for articulating this

concept was to be accused of sentimentality unbecoming a judge as well as seeking to make litigation the ultimate recourse for most harms (McCall 1984: 745; Ursin 2009: 1291). His own view was more nuanced – he acknowledged that cost shifting was an element of his proposal, but far from the only one (Traynor 1965:366-369). At the same time, Traynor openly acknowledged product liability was a question of policy rather than a timeless principle of law. This may have had something to do with the fact that Traynor held a doctorate in political science - to him, the law could not stand outside of time as formalists argued, or rest on some set of “neutral principles” articulated by the legal process school (McCall 1984: 743- 745; Ursin 2009: 1320-1324).

Like any true policy entrepreneur, Traynor used many different opportunities and channels to advance his position until eventually the view of the rest of the court caught up to him. In *Greenman v. Yuba Power Products, Inc* (Cal. 1963), 377 P.2d 897, Traynor wrote an opinion embracing strict liability not just for a majority but a unanimous bench. He followed this with a course of advocacy in law reviews and public speaking during and after his tenure on the bench (Traynor 1965; McCall 1984: 745-746). From Traynor’s work grew entire bodies of law relating to product liability and enterprise liability that govern many modern suits and a *Restatement of Law* is dedicated to the concept. Thus, in the span of a few decades a Kuhnian revolution in law on this point was nearly complete, with Traynor having moved from heretic to beatified founder.

Whether there is someone willing and able to serve as the policy entrepreneur on behalf of more rigorous, in depth scrutiny of expert witnesses is a doubtful proposition. The *Reference Manual on Scientific Evidence* suggests some advocates both in the law

and scientific fields care passionately about trying to ensure that legal judgments reflect accurate scientific information. However, it stops short of either embracing the full range of testimony judges confront or advocating for a more active role for judges in assessing expert testimony. Its infrequent citation in this study's cases speaks to the fact that at present it is no substitute for a shift in norms or stricter guidance.

In the absence of such a change the gatekeeping function of *Daubert/Kumho* will likely continue in the pattern identified in this study. Judges will follow the law by examining issues surrounding expert witness testimony. They will also, in an ironic echo of Holmes' ([1897]1997: 993-994) "bad man," do only what compliance requires in most cases, especially where engaging the material would require high cognitive effort or transgress their views of appropriate judicial conduct. More generally, legal policy makers would do well to consider how any new legal policy – in the form of a ruling or a piece of legislation – comports with those normative views and seek to ensure alignment or minimize discretion as appropriate. Policy entrepreneurs could also take a lesson from the Federalist Society and seek to reshape norms through the institution of legal training and socialization. Taking such direct action is transgressive to many legal actors, but a failure to do so both cedes the field to persons lacking such constraints and all but ensures sub-optimal execution of what might be perfectly sound legal policy.

D. Practitioners can Maximize the Probabilities of Success on Challenges to Expert Witnesses by Emphasizing Norm Violations Rather than Technical Arguments

As a purely practical matter, this study also has insights for attorneys considering making an objection to the testimony of an expert witness. The first is that successful motions to exclude typically succeed for the same reason any other motion does, rather

than applying the unique factors of *Daubert*. The odds of exclusion might also be enhanced if the consequence is that the judge can resolve the case. Similarly, getting an expert's testimony restricted is often easiest if the argument aligns with judicial concerns about wasted effort normative violations rather than asking the judge to wade into murky methodological questions.

One pattern that was striking in the analysis of these cases was the faith judges have in the adversarial system. While an advocate might have pointed out what appeared like a glaringly obvious error, the typical response was to state it went to the weight rather than the admissibility of the evidence. Judges firmly believe in the power of cross-examination and adversarial presentations, which is a natural outgrowth of their role and training. They are unlikely to be moved by social science studies noting the problems with using these processes to identify and counter problematic testimony, even if they do not join the Chief Justice in labeling it “gobbledygook.” (Rocco 2017).

In light of that tendency, advocates seeking to exclude an expert would do well to avoid invoking the *Daubert* factors unless the testimony fits what a judge sees as “scientific” rather than experiential or technical expertise. “Scientific” experts tend to come from fields such as biology, toxicology, chemistry and occasionally even sociology, but generally not architecture, clinical psychology, engineering, accounting, property evaluation, medicine, vocational counseling, fire investigation or even most forms of accident reconstruction. Even if the evidence falls into this restrictive circle, the problem should be glaringly obvious such as a disregard for established protocols or a complete lack of evidence of replicability/testability, error rate, peer review and general

acceptance. Anything short of this is apt to be assessed as an issue of weight for the jury to decide.

This does not mean testimony outside of the fields of judicially constructed “science” will always be admitted. It is just that success on this sort of motion typically requires proof of an extreme flaw that renders the testimony so suspect or irrelevant that it would waste the jury’s time to even hear it. This is not as difficult as it may sound – as with most litigation, the most powerful evidence was typically the testimony of the challenged expert either in deposition or a *Daubert* hearing. In other words, just because *Daubert/Kumho* supplies the rule of law does not mean this motion is any different than most other elements of legal practice.

In a very real sense this speaks to the power of cognitive simplification – judges generally respond to the sort of evidence that would sway them in any other setting. Judges are used to assessing testimony and using admissions against interest against a party. They are also used to comparing evidence to a pre-existing set of elements and deciding a key point has been omitted. Thus, an expert who admits their conclusions are speculative, even if it is meant rhetorically, is easy for the judge to eliminate. Similarly, when the expert admits to deviating from an established standard without justification or having no experience in a field it creates a much simpler task than if they have to wade through *Daubert/Kumho*’s specific list of factors. Indeed, invoking the factors without a more compelling narrative or piece of testimony appears unlikely to move a judge to action. Hence, both those seeking to admit an expert and those looking to exclude one would do well to think of how a judge would evaluate the witness in any other setting

before they attempt to rest their attack or defense on the complexities of testability, error rate, peer review and general acceptance.

Appealing to a judge's existing preferences and interests indirectly may also be a way to enhance the chances of excluding an expert witness. As noted above, co-presenting a motion for exclusion with a motion for summary judgment was associated with an elevated rate of exclusions. Although this relationship did not demonstrate statistical significance in the logistic regression model for exclusions, it was persistent and universal in all of the examined court clusters. It may be that if the expert can be presented as the key to being able to resolve the case it may create a strategic incentive for the court to grant the motion, whereas in a case which is going to go to trial either way no such incentive exists. Even here, success requires a clear issue with the testimony – otherwise not only is rejection unlikely but even a restriction on the testimony may not be granted on the idea that such a ruling would be premature.

More generally, litigants should be realistic about the likely outcome of their motion. Exclusions are rare and curbing an expert's testimony is more likely than excluding it. Moreover, restrictions, while more common, are even less likely to be predicated on a challenge to the methods employed by the expert rather than the manner of their testimony. Judges are far more responsive to allegations that the expert is going into either irrelevant testimony or areas reserved for the judge or jury. Many of the restrictions granted, for example were to bar testimony that instructed the jury on who to believe or what the legal standards are. While this might not sound like much of an impact, it prevents the expert from giving the jury a simple, direct statement of the

party's argument from the witness stand – which is often the main point of their testimony.

E. Judges are, in the end, Human and Complex – even when Deciding Whether to Admit Expert Witnesses

Going back to the start of this dissertation, the answer to the question of what influences district court decision making in the context of admitting expert witnesses is fairly clear. The central influence, the one that drives the outcome, is precisely what judges say it is – the legal command in *Daubert* and *Kumho* to default in favor of the admission of expert testimony and reserve most issues of whether to accept the expert's testimony as true for the jury. This is true in every court and with every group of judges examined – they default to admission and favor permitting even flawed testimony to intervening on methods. Although the relative rate of admissions varies by location and whether a judge is appointed by a Democratic or a Republican President, the overwhelming tendency is to admit rather than exclude. Moreover, the rate showed no consistent pattern in favor or against any particular type of party. This in turn justifies concluding judges are following the rule of law in these settings as they see it, and likely are doing so with their other pre-trial decisions in most cases as well.

At the same time, the manner that judges act in applying the *Daubert/Kumho* standards suggest that in following the command to default to admission judges may be sacrificing some of the rigor of the gatekeeping process envisioned by the Court. Judges are quite willing to restrict testimony that they deem irrelevant or outside of the proper scope of expert testimony, but do not seem inclined to investigate the methods of experts

in detail in most cases. Most method concerns, outside of extreme cases, are reserved for the jury to sort out in keeping with the norms and practices of adversarial legalism.

As a body, judges seem to be doing what they can to minimize the cognitive burden of this analysis. They delay motions where they deem them unnecessary. They craft boilerplate analyses of the law and sometimes the facts and apply them as a time and effort saving measure that limits any value of the writing process as a counter to heuristic tendencies. They have simplified the process of resolving these disputes to a handful of steps, and essentially jettisoned the *Daubert/Kumho* factors of testing, error rate, peer review and general acceptance outside of a limited scope of cases. Instead, an examination of the expert's methods frequently reduces to an assurance that the expert is following a socially constructed standard deriving from the actions of other courts or industry groups, without any apparent assurance that those standards generate truly reliable testimony.

These are completely understandable steps given the burden district courts are under, the way *Daubert/Kumho* structures the choices and options for them and the engrained norms of and expectations of the judicial role within adversarial legalism. It speaks to the humanity of judges that they are cognitive misers, even when they put on their robes. If the law were more restrictive, the limited influence of location and partisanship might be even more remote, and less evidence of cognitive shortcuts might be present. Without the clear indicators in *Daubert/Kumho* to default to admission, other factors might well become more important. In other words, as with all other human

endeavors, district court judicial decision making is messy, nuanced and contextually dependent.

To truly investigate judicial decision making, scholars are going to have to be willing to grapple with this complexity. Similarly, to effectuate significant policy change policy makers, whether they are judges or legislators, will have to consider how a proposed change comports with the norms of judges, lawyers and society – they cannot simply assume compliance will be robust because of their position within a hierarchical system. Legal practitioners should try to remember that the judge deciding their *Daubert/Kumho* motion is not an automaton applying the pronouncements of the Supreme Court as they learned them in law school but a harried public servant trying to handle an overburdened docket and clear cases with minimal effort and tailor their efforts accordingly. If scholars, policy makers and the legal community could acknowledge the humanity of judges and their common commitment to the rule of law as they see it through their differentiated experience, it may even be able to abate the divisive politics over judicial appointments and the President's ill-informed smear of Judge Curiel would not have a refuge in the more outlandish pronouncements of political scientists about judges as politicians in robes.

References

Books, articles and statutes

- Bailey, Michael A. and Maltzman, Forrest. 2011. *The Constrained Court: Law, Politics and the Decisions Justices Make*. Princeton, NJ: Princeton University Press.
- Bainbridge, Stephen M., and Gulati, Mitu. 2002. "How Do Judges Maximize? (The Same Way Everybody Else Does-Boundedly): Rules of Thumb in Securities Fraud Opinions." *Emory Law Journal*, 51:83-151.
- Baker, Tom. 2005. "Reconsidering the Harvard Medical Practice Study Conclusions about the Validity of Medical Malpractice Claims." *The Journal of Law, Medicine and Ethics*, 33:501-514.
- Baum, Lawrence. 2006. *Judges and Their Audiences: A Perspective on Judicial Behavior*. Princeton, NJ: Princeton University Press.
- Baum, Lawrence. 2010. "Motivation and Judicial Behavior: Expanding the Scope of Inquiry." In *The Psychology of Judicial Decision Making*, ed. David E. Klein and Gregory Mitchell, 1-26. Oxford, UK: Oxford University Press.
- Beebe, Barton. 2006. "An Empirical Study of the Multifactor Tests for Trademark Infringement." *California Law Review*, 94:1581-1633.
- Beecher-Monas, Erica. 2001. "The Heuristics of Intellectual Due Process: A Primer for Triers of Science." *NYU Law Review*, 75:1564-1657.
- Beecher-Monas, Erica. 2011. "Expert Testimony in Civil Cases." In *The Future of Evidence: How Science and Technology Will Change the Practice of Law*, ed. Carol Henderson and Jules Epstein, 33-76. Chicago: American Bar Association.
- Berger, Linda L. 2013. "A Revised View of the Judicial Hunch." *J. ALWD*, 10:1-39.
- Beth, Richard S. 2013. "Cloture Attempts on Nominations: Data and Historical Development." Congressional Research Service.
- Beyea, Jan, and Berger Daniel. 2001. "Scientific Misconceptions among *Daubert* Gatekeepers: The Need for Reform of Expert Review Procedures." *Law and Contemporary Problems*, 64(2/3):327-372.
- Bhatia, Kedar. 2016. "Final October Term 2015 Stat Pack." Available at <http://www.scotusblog.com/2016/06/final-october-term-2015-stat-pack/>. Accessed February 12, 2018.

- Blackstone, William. 1765. *Commentaries on the Laws of England*. New Haven, Conn: The Avalon Project.
- Bornstein, Brian H. and Robicheaux, Timothy R. 2008. "Crisis, What Crisis? Perception and Reality in Civil Justice." In *Civil Juries and Civil Justice: Psychological and Legal Perspectives*, ed. Brian H. Bornstein, Richard L. Wiener, Robert Schopp, and Steven L. Willborn, 1-22. New York: Springer.
- Boyd, Christina L. 2013. "She'll Settle It?" *Journal of Law and Courts*, 1(6):193-219.
- Boyd, Christina L. 2016. "Representation on the Courts? The Effects of Trial Judges' Sex and Race." *Polit. Res. Q.* 69:788-99.
- Braun, Virginia and Clarke, Victoria. 2006. "Using Thematic Analysis in Psychology." *Qualitative Research in Psychology*. 3:77-101.
- Brekke, Nancy J.; Enko, Peter J.; Clavet, Gail; Seelau, Eric. 1991. "Of Juries and Court-Appointed Experts: The Impact of Nonadversarial versus Adversarial Expert Testimony." *Law and Human Behavior*, 15:451-475.
- Breyer, Stephen. 2010. *Making Our Democracy Work: A Judge's View*. New York: Alfred A. Knopf.
- Brown, Karina. 2017. "9th Circuit Nominee Could Bolster Trump's Immigration Policy" *Courthouse News Service*. Available at <https://www.courthousenews.com/9th-circuit-nominee-bolster-trumps-immigration-policy/>. Accessed on January 19, 2018.
- Buchman, Jeremy. 2007. "The Effects of Ideology on Federal District Judges' Decisions to Admit Expert Testimony." *American Politics Research*, 35:671-693.
- Carp, Robert A. and Rowland, Charles K. 1983. *Policymaking and Politics in the Federal District Courts*. Lawrence, KS: University of Kansas Press.
- Carter, Lief H., and Burke Thomas F. 8th ed. 2010. *Reason in Law*. New York: Longman.
- Cecil, Joe S. and Willging, Thomas E. 1993. "Court-Appointed Experts: Defining the Role of Experts Appointed Under Federal Rule of Evidence 706". Federal Judicial Center.
- Charmaz, Kathy. 2006. *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis*. Thousand Oaks: Sage Publications.

- Cheng, Edward K. and Yoon, Albert H. 2005. "Does *Frye* or *Daubert* Matter? A Study of Scientific Admissibility Standards." *Virginia Law Review*, 91:471-513.
- Clayton, Cornell W. 1999. "The Supreme Court and Political Jurisprudence: New and Old Institutionalisms". In *Supreme Court Decision Making: New Institutional Approaches*, ed. Cornell W. Clayton and Howard Gillman, 15-42. Chicago, IL: University of Chicago Press.
- Clayton, Cornell W., and Gillman, Howard eds. 1999. *Supreme Court Decision Making: New Institutional Approaches*. Chicago, IL: University of Chicago Press.
- Collins, Michael. 2016. "Trump Victory is the end of the line for Obama judicial nominees." *USA Today*. Available at www.usatoday.com/story/news/politics/2016/11/11/trump-victory-end-line-obama-judicial-nominees/93615006/. Accessed on Aug 3, 2017.
- Cooper, Philip J. 1988. *Hard Judicial Choices: Federal District Court Judges and State and Local Officials*. New York: Oxford University Press.
- Cooper, Joel and Neuhaus, Isaac M. 2000. "The "Hired Gun" Effect: Assessing the Effect of Pay, Frequency of Testifying, and Credentials on the Perception of Expert Testimony." *Law and Human Behavior*, 24:149-171.
- Cwik, Cynthia, Garrett, N.P., O'Donnell, K.V., and Sanberg, M.A. 2013. "Expert Evidence in the Ninth Circuit." In *Scientific Evidence Review: Admissibility and Use of Expert Evidence in the Courtroom*, ed. Cynthia Cwik, Jules Epstein, and Carol Henderson, Monograph No. 9:323-365. Chicago: American Bar Association.
- Dahl, Robert A. 1957. "Decision making in a Democracy: The Supreme Court as a National Policy-Maker." *Journal of Public Law*, 5:279-295.
- Diamond, Jared. 2012. *The World Until Yesterday: What Can We Learn from Traditional Societies*. New York: Penguin Books.
- Dixon, Lloyd and Gill, Brian. 2002. "Changes in the Standards for Admitting Expert Evidence in Federal Civil Cases Since the *Daubert* Decision." *Psychology, Public Policy, and Law*, 8:251-308.
- Dworkin, Ronald. 1977. *Taking Rights Seriously*. Cambridge, MA: Harvard University Press.
- Dworkin, Ronald. 1986. *Law's Empire*. Cambridge, MA: The Belknap Press of Harvard University Press.

- Eder, Steve and Medina, Jennifer. 2017. "Trump University Settlement Approved by Judge". *New York Times*. Available at www.nytimes.com/2017/03/32/us/trump-university-settlement.html. Accessed on Aug. 3, 2017.
- Edwards, Harry T. and Livermore, Michael A. 2009. "Pitfalls of Empirical Studies that Attempt to Understand the Factors Affecting Appellate Decisionmaking." *Duke Law Journal*, 58:1895-1989.
- Egelko, Bob. 2014. "Judge Vaughn Walker Tells his Side of Prop. 8 Trial". *SF Gate*. Available at <http://www.sfgate.com/lgbt/article/Judge-Vaughn-Walker-tells-his-side-of-Prop-8-5416851.php>. Accessed on Aug 3, 2017.
- Epstein, Lee. 2016. "Some Thoughts on the Study of Judicial Behavior." *Wm. & Mary Law Rev.* 57:2017-2073.
- Epstein, Lee, and Knight, Jack. 1998. *The Choices Justices Make*. Washington, DC: CQ Press.
- Epstein, Lee, and Knight, Jack. 2013. "Reconsidering judicial preferences." *Annual Review of Political Science*. 16:11-31.
- Epstein, Lee and Walker, Thomas. 8th ed. 2013. *Constitutional Law for a Changing America: Institutional Powers and Constraints*, Washington, DC: CQ Press.
- Epstein, William. 1981. "The Historical Foundation of the Methodology of American Legal Education." *The Journal of Legal History*, 2:3-29.
- Federal Bureau of Investigation (2015). "FBI Testimony on Microscopic Hair Analysis Contained Errors in at Least 90 Percent of Cases in Ongoing Review". Press Release. Available at <https://www.fbi.gov/news/pressrel/press-releases/fbi-testimony-on-microscopic-hair-analysis-contained-errors-in-at-least-90-percent-of-cases-in-ongoing-review>. Accessed on January 16, 2018.
- Federal Judicial Center, National Research Council of the National Academies. 3rd ed. 2011. *Reference Manual on Scientific Evidence*. Washington, D.C.: The National Academies Press.
- Fessenden, Helen. 2014. "Suzanne Bonamici Shares Lessons from 2 Years in the U.S. House" *The Oregonian/Oregon Live*. Available at http://www.oregonlive.com/politics/index.ssf/2014/02/suzanne_bonamici_shares_lesson.html. Accessed on January 19, 2018.
- Feeley, Malcolm M. and Rubin, Edward L. 1998. *Judicial Policy Making and the Modern State: How the Courts Reformed America's Prisons*. Cambridge, UK: Cambridge University Press.

- Frank, Jerome. 4th printing 1936. *Law and the Modern Mind*. New York: Coward-McCann.
- Fuller, Lon L. Rev. ed. 1977. *The Morality of Law*. New Haven, CT: Yale University Press.
- Galanter, Marc. 1974. "Why the 'Haves' Come Out Ahead: Speculations on the Limits of Legal Change." *Journal of Law and Society*, Fall 1974:95-160.
- Galanter, Marc. 1993. "The Regulatory Function of the Civil Jury." In *Verdict: Assessing the Civil Jury System*, ed. Robert E. Litan, 61-85. Washington, DC: Brookings Institution.
- Gallup. 2017. "Supreme Court Job Approval." Available at <http://www.gallup.com/poll/4732/supreme-court.aspx> . Accessed on Aug 3, 2017.
- Gastil, John; Deess, E. Pierre; Weiser, Philip J.; and Simmons, Cindy. 2010. *The Jury and Democracy: How Jury Deliberation Promotes Civic Engagement and Political Participation*. Oxford, UK: Oxford University Press.
- Gatowski, Sophia I.; Dobbin, Shirley A.; Richardson, James T.; Ginsburg, Gerald P.; Merlino, Mara L.; and Dahir, Veronica. 2001. "Asking the Gatekeepers: A National Survey of Judges on Judging Expert Evidence in a Post-*Daubert* World." *Journal of Law and Human Behavior*, 25:433-458.
- Geyh, Charles Gardner. 2016. *Courting Peril: The Political Transformation of the American Judiciary*. New York, NY: Oxford University Press.
- Giles, Micheal W.; Hettinger, Virginia A.; and Peppers, Todd. 2001. "Picking Federal Judges: A Note on Policy and Partisan Selection Agendas." *Political Research Quarterly*, 54:623-64.
- Guthrie, Chris; Rachlinski, Jeffrey J.; and Wistrich, Andrew J. 2007. "Blinking on the Bench: How Judges Decide Cases." *Cornell Law Review*, 93:1-43.
- Hamilton, Alexander. [1788]2003. "Federalist No. 78." In *The Federalist Papers*, Alexander Hamilton, James Madison and John Jay, ed. Clinton Rossiter, 463-471. New York: Signet Classic.
- Hart, H.L.A. 2nd ed. 1994. *The Concept of Law*. Oxford: Oxford University Press.
- Heise, Michael. 2002. "The Past, Present, and Future of Empirical Legal Scholarship: Judicial Decision Making and the New Empiricism." *U. Ill. L. Rev.* 2002(4):819-850.

- Helland, Eric and Klick, Jonathan. 2012. "Does Anyone Get Stopped at the Gate? An Empirical Assessment of the *Daubert* Trilogy in the States." *Supreme Court Economic Review*, 20:1-33.
- Holmes, Jr., Oliver Wendell. 1881. *The Common Law*. Boston: Little, Brown and Company.
- Holmes, Jr., Oliver Wendell. [1897] 1997. "The Path of the Law." Reprinted in *Harvard Law Review*, 110:991-1009.
- Hornby, D. Brock. [2007] 4th ed. 2013. "The Business of US District Courts." In *Judges on Judging*, David M. O'Brien, 106-114. Thousand Oaks, CA: CQ Press.
- Horwitz, Paul. 2013. "What Ails the Law Schools." *Michigan Law Review*, 111:955-976.
- Hsieh, Hsiu-Fang and Shannon, Sarah E. 2005. "Three Approaches to Qualitative Content Analysis." *Qualitative Health Research*, 15:1277-1288.
- Jasanoff, Sheila. 1995. *Science at the Bar: Law, Science, and Technology in America*. Cambridge, MA: Harvard University Press.
- Jasanoff, Sheila. 2008. "Representation and Re-Presentation in Litigation Science." *Environmental Health Perspectives*, 116:123-129.
- Johnson, R. Burke; Onwuegbuzie, Anthony J.; and Turner, Lisa A. 2007. "Towards a Definition of Mixed Methods Research." *Journal of Mixed Methods Research*, 1:112-133.
- Jonakait, Randolph N. 2003. *The American Jury System*. New Haven, CT: Yale University Press.
- Jones, Jeffrey M. (2015). "Trust in US Judicial Branch Sinks to New Low of 53%." Gallup. Available at http://www.gallup.com/poll/185528/trust-judicial-branch-sinks-new-low.aspx?g_source=position4&g_medium=related&g_campaign=tiles Accessed on Aug 3, 2017.
- Justice, William W. [1992] 4th ed. 2013. "The Two Faces of Judicial Activism." In *Judges on Judging*, David M. O'Brien, 42-55. Thousand Oaks, CA: CQ Press.
- Kagan, Robert A. 2001. *Adversarial Legalism: The American Way of Law*. Cambridge, MA: Harvard University Press.
- Kahan, Dan M.; Hoffman, David A.; Evans, Danielie; Devins, Neal; Lucci, Eugene A.; and Cheng, Katherine. 2016. "'Ideology' or 'Situation Sense'? An Experimental

- Investigation of Motivated Reasoning and Professional Judgment.” *University of Pennsylvania Law Review*, 164:349-439.
- Kahneman, Daniel. 2011. *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux.
- Kahneman, Daniel and Tversky, Amos 1979. “Prospect Theory: An Analysis of Decision Under Risk.” *Econometrica*, 47(2):263-291.
- Keck, Thomas. 2014. *Judicial Politics in Polarized Times*. Chicago: University of Chicago Press.
- Kim, Pauline T.; Schlanger, Margo; Boyd, Christina L.; and Martin, Andrew D. 2009. “How Should We Study District Judge Decision making.” *Washington Univ. Journal of Law and Policy*, 29:83-112.
- Kozinski, Alex [1993] 4th ed. 2013. “What I ate for Breakfast and Other Mysteries of Judicial Decision Making.” In *Judges on Judging*, David M. O’Brien, 115-120. Thousand Oaks, CA: CQ Press.
- Kritzer, Herbert M. 2007. “Law is the Mere Continuation of Politics by Different Means: American Judicial Selection in the Twenty-First Century.” *DePaul Law Review* 56:423-467.
- Kuersten, Ashlyn and Songer, Donald. 2014. *Decisions on the US Courts of Appeals*. New York: Routledge.
- Kuhn, Thomas. 3rd ed. 1996. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.
- Levett, Lora M. and Kovera, Margaret B. 2008. “The Effectiveness of Opposing Expert Witnesses for Educating Jurors About Unreliable Expert Evidence.” *Law and Human Behavior*, 32:363-374.
- Levitan, Daniel J. 2017. *Weaponized Lies- How to Think Critically in the Post-Truth Era*. New York: Dutton.
- Lind, E. Allan and Tyler, Tom R. 1988. *The Social Psychology of Procedural Justice*. New York: Plenum Press.
- Liptak, Adam. 2017. “An exit Interview with Richard Posner, Judicial Provocateur.” *The New York Times*. Available at www.nytimes.com/2017/09/11/us/politics/judge-richard-posner-retirement.html. Accessed November 13, 2017.

- Locke, John [1690]2002. *The Second Treatise of Government and A Letter Concerning Toleration*. Mineola, NY: Dover Publications, Inc.
- Marcus, Ruth. 2015. "The Judge Immigration Foes Wanted." *The Washington Post*. Available at www.washingtonpost.com/opinions/the-judge-immigration-foes-wanted/2015/02/17/9afb0de-b6e8-11e4-aa05-1ce812b3fdd2_story.html?utm_term=.2e9b322fb6d4 Accessed on Aug 3, 2017.
- McAuliff, Bradley D. and Duckworth, Tejah D. (2010). "I Spy With My Little Eye: Jurors' Detection of Internal Validity Threats in Expert Evidence." *Law and Human Behavior*, 34:489-500.
- McAuliff, Bradley D.; Kovera, Margaret B.; and Nunez, Gabriel. 2009. "Can Jurors Recognize Missing Control Groups, Confounds, and Experimenter Bias in Psychological Science?" *Law and Human Behavior*, 33:247-257.
- McCall, James R. 1984. "Roger Traynor: Teacher, Jurist and Friend." *Hastings L.J.* 35:741-746.
- Merlino, Mara L.; Springer, Victoria; and Sigillo, Alexandra E. 2011. "The Social Construction of the Admissibility of the Most Frequently Proffered Varieties of Expert Testimony. In *The Future of Evidence: How Science and Technology Will Change the Practice of Law*, ed. Carol Henderson and Jules Epstein, 1-32. Chicago: American Bar Association.
- Miles, Thomas J. and Sunstein, Cass R. 2008. "The New Legal Realism." *The University of Chicago Law Review*, 75:831-851.
- Morgan, Douglas F. and Shinn, Craig W. 2014. "The Foundations of New Public Governance." In *New Public Governance: A Regime-Centered Perspective*, ed. Douglas F. Morgan and Brian J. Cook, 3-12. Armonk, NY: M.E. Sharpe.
- Neubauer, David W. and Meinhold, Stephen S. 6th ed. 2013. *Judicial Process: Law, Courts and Politics in the United States*. Boston, MA: Wadsworth.
- O'Brien, David M. 1999. "Institutional Norms and Supreme Court Opinions: On Reconsidering the Rise of Individual Opinions." In *Supreme Court Decision Making: New Institutional Approaches*, ed. Cornell W. Clayton and Howard Gillman, 91-114. Chicago: University of Chicago Press.
- Odendahl, Marilyn. 2012. "Southern District Judiciary Comes Together at 'Historic' Event to Thank Lugar, Bayh" *The Indiana Lawyer*. Available at <https://www.theindianalawyer.com/articles/29974-southern-district-judiciary-comes-together-at-historic-event-to-thank-lugar-bayh>. Accessed on January 20, 2018.

- Oregonian Editorial Board. 2011. "What a Long, Strange Trip to the Federal Bench" *The Oregonian/OregonLive*. Available at http://www.oregonlive.com/opinion/index.ssf/2011/02/what_a_long_strange_trip_to_th.html. Accessed January 19, 2018.
- Pew Research Center. 2014. "Political Polarization in the American Public." Available at www.people-press.org/2014/06/12/political-polarization-in-the-american-public. Accessed Aug. 3, 2017.
- Pollock III, Philip H. 4th Edition 2011. *The Essentials of Political Analysis*. New York: CQ Press.
- Posner, Richard A. 1995. *Overcoming Law*. Cambridge, MA: Harvard University Press.
- Posner, Richard A. 2008. *How Judges Think*. Cambridge, MA: Harvard University Press.
- Rachlinski, Jeffrey J.; Guthrie, Chris; and Wistrich, Andrew J. 2013. "Altering Attention in Adjudication." *UCLA Law Review*, 60:1586-1618.
- Rachlinski, Jeffrey and Wistrich, Andrew. 2017. "Judging the Judiciary by the Numbers: Empirical Research on Judges." *Ann. Rev. Law Soc. Sci.* 13:203-229.
- Robbennolt, Jennifer K. MacCoun, Robert J.; Darley, John M. 2010. "Multiple Constraint Satisfaction in Judging". In *The Psychology of Judicial Decision Making*, ed. David E. Klein and Gregory Mitchell, Gregory, 27-40. Oxford, UK: Oxford University Press.
- Robbins, Ted. 2017. "Who is Judge James L. Robart and why did he block Trump's Immigration Order?" *NPR*. Available at <http://www.npr.org/2017/02/04/513446463/who-is-judge-james-l-robart-and-why-did-he-block-trumps-immigration-order>. Accessed Aug 3, 2017.
- Rocco, Phillip. 2017. "Justice Roberts said Political Science is 'Sociological Gobbledygook.' Here's why he said it, and Here's why he's Mistaken." *The Washington Post*. Available at www.washingtonpost.com/news/monkey-cage/wp/2017/10/04/justice-roberts-said-political-science-is-sociological-gobbledygook-heres-why-he-said-it-and-why-hes-mistaken/?utm_term=.96ab8de54bfd. Accessed Nov. 13, 2017.
- Rossoll, Nicki. 2016. "Donald Trump Calls Chief Justice John Roberts a 'Nightmare for Conservatives'" *ABC News*. Available at <http://abcnews.go.com/Politics/donald-trump-calls-chief-justice-john-roberts-nightmare/story?id=36336627>. Accessed on Aug. 3, 2017.

- Rowland, Charles K. and Carp, Robert A. 1996. *Politics and Judgment in Federal District Court*. Lawrence, KS: University of Kansas Press.
- Rowland, Charles K.; Traficanti, Tina; and Vernon, Erin. 2010. "Every Jury Trial is a Bench Trial: Judicial Engineering of Jury Disputes." In *The Psychology of Judicial Decision Making*, ed. David E. Klein and Gregory Mitchell, 183-202. Oxford, UK: Oxford University Press.
- Schauer, Frederick, and Spellman, Barbara A. 2013. "Is Expert Evidence Really Different." *Notre Dame L. Rev.*, 89:1-26.
- Schwartz, Victor E., and Silverman, Cary. 2006. "The Draining of *Daubert* and the Recidivism of Junk Science in Federal and State Courts." *Hofstra L. Rev.* 35:217-273.
- Seawright, Jason and Gerring, John. 2008. "Case Selection Techniques in Case Study Research: A Menu of Qualitative and Quantitative Options." *Political Research Quarterly*, 61:294-308.
- Segal, Jeffrey A. and Spaeth, Harold J. 2002. *The Supreme Court and the Attitudinal Model Revisited*. Cambridge, UK: Cambridge University Press.
- Sevier, Justin. 2014. "The Truth-Justice Tradeoff: Perceptions of Decisional Accuracy and Procedural Justice in Adversarial and Inquisitorial Legal Systems." *Psychology, Public Policy, and Law*, 20:212-224.
- Shapiro, Martin. 1965. "Stability and Change in Judicial Decision making: Incrementalism or Stare Decisis." *Law in Transition Quarterly*, 2:134-157.
- Sherman, Mark. 2016. "Where do the presidential candidates stand on Supreme Court nominees?" *PBS Newshour*. Available at www.pbs.org/newshour/rundown/presidential-candidates-stand-supreme-court-nominees/. Accessed on Aug 3, 2017.
- Simon, Herbert. 1955. "A Behavioral Model of Rational Choice." *Quarterly Journal of Economics*, 69:99-118.
- Slotnick, Elliot; Schiavoni, Sara; and Goldman, Sheldon 2017. "Obama's Judicial Legacy: The Final Chapter." *Journal of Law and Courts*, 5:363-422.
- Songer, Donald R., Segal, Jeffrey A., and Cameron, Charles M. 1994. "The Hierarchy of Justice: Testing a Principal-Agent Model of Supreme Court-Circuit Court Interactions." *American Journal of Political Science*, 38:673-696.

- Songer, Donald R., Sheehan, Reginald S., and Haire, Susan B. 2003. *Continuity and Change on the United States Court of Appeals*. Ann Arbor, MI: The University of Michigan Press.
- Taha, Ahmed E. 2004. "Publish or Paris? Evidence of How Judges Allocate Their Time." *American Law and Economics Review*, 6:1-27.
- Tamanaha, Brian Z. 2009. "The Distorting Slant in Quantitative Studies of Judging," *BC Law Review*, 50:685-758.
- Thaler, Richard H. and Sunstein, Cass R. 2008. *Nudge: Improving Decisions About Health, Wealth, and Happiness*. New Haven, CT: Yale University Press.
- Totenberg, Nina. 2016. "Trump Presses Case That 'Mexican' Judge Curiel is Biased Against Him." *NPR*. Available at www.npr.org/2016/06/04/580714972/trump-presses-case-that-mexican-judge-curiel-is-biased-against-him. Accessed on Aug. 3, 2017.
- Traynor, Roger J. 1965. "The Ways and Meanings of Defective Products and Strict Liability." *Tennessee Law Review*, 32:363-376.
- Tversky, Amos and Kahneman, Daniel. 1974. "Judgment under Uncertainty: Heuristics and Biases." *Science*, 185:1124-1131.
- United States District Court, District of Oregon. 2017. "Role of Magistrate Judges in the District of Oregon". Available at www.ord.uscourts.gov/index.php/court-info/court-policies/role-of-magistrates. Accessed on November 17, 2017.
- Ursin, Edmund. 2009. "How Great Judges Think: Judges Richard Posner, Henry Friendly and Roger Traynor on Judicial Lawmaking." *Buffalo Law Review*, 57:1267-1360.
- Vidmar, Neil. 1995. *Medical Malpractice and the American Jury: Confronting the Myths About Jury Incompetence, Deep Pockets, and Outrageous Damage Awards*. Ann Arbor, MI: University of Michigan Press.
- Watson, Randy. 2015. "Is it Arson? The NFPA 921 Effect". *In-House Defense Quarterly*, Spring 2015:24-27, 84.
- Wechsler, Herbert 1959. "Toward Neutral Principles of Constitutional Law." *Harvard Law Rev.*, 73:1-35.
- Wheeler, Lydia. 2015. "Faith leaders demand that liberal justices sit out gay marriage case." *The Hill*. Available at <http://thehill.com/regulation/240163-religious-leaders-want-justices-restrained-from-ruling-on-same-sex-marriage>. Accessed on Aug. 3, 2017.

Williams, Joseph P. 2017. "The Battle for the Judiciary." *US News and World Reports*. Available at www.usnews.com/news/the-report/articles/2017-08-02/democrats-take-aim-at-trumps-judicial-appointments. Accessed on Aug. 3, 2017.

Wistrich, Andrew; Guthrie, Chris; and Rachlinski, Jeffrey J. 2005. "Can Judges Ignore Inadmissible Information? The Difficulty of Deliberately Disregarding." *U. Penn Law Rev.*, 153:1251- 1345.

Wong, John B., Gostin, Lawrence O. and Cabrera, Oscar A. 2011. "Reference Guide on Medical Testimony." In *Reference Manual on Scientific Evidence*, 3rd ed., Federal Judicial Center, National Research Council of the National Academies, 687-746. Washington, D.C.: The National Academies Press.

Wyden, Ron and Merkley, Jeffrey. 2017. "Letter to Donald F. McGahn II." Available at <https://www.almcms.com/contrib/content/uploads/sites/292/2017/09/WH-Judicial-Vacancy-Signed.pdf>. Accessed on January 19, 2018.

Zaring, David. 2004. "National Rulemaking Through Trial Courts: The Big Case and Institutional Reform." *UCLA Law Review*, 51:1015-1078.

Cases

3600 Michigan Co., Ltd. v. Infra-Metals, Co. N.D. Ill. 2011. 2011 U.S. Dist. LEXIS 574.

Agrigenetics v. Pioneer Hi-Bred International, Inc. S.D. Ind. 2010. 2010 U.S. Dist. LEXIS 12003.

Apilad v. North American Gay Amateur Athlete Alliance. W.D. Wash. 2011. 2011 U.S. Dist. LEXIS 159575.

Aurand v. Norfolk Southern Railway Co. N.D. Ind. 2011. 802 F.Supp. 2d 950.

Bingham v. Raytheon Technical Services Co., LLC. S.D. Ind. 2014. 2014 U.S. Dist. LEXIS 160499.

Bisson v. BNSF Railway Co. E.D. Wash. 2015. 2015 U.S. Dist. LEXIS 182242.

Bryant v. Wyeth. W.D. Wash. 2012. 2012 U.S. Dist. LEXIS 190912.

C.A. v. AMLI at Riverbend, L.P. S.D. Ind. 2010. 2010 U.S. Dist. LEXIS 86530.

Cascade Yarns, Inc. v. Knitting Fever, Inc. W.D. Wash. 2012. 2012 U.S. Dist. 15097.

Citizens Ins. Co. of the Midwest v. LG Electronics USA, Inc. S.D. Ind. 2012. 2012 U.S. Dist. LEXIS 1127128.

C.W. v. Textron, Inc. 7th Cir. 2015. 807 F.3d 827.

Dasho v. City of Federal Way. W.D. Wash. 2015. 101 F. Supp. 3d 1025.

Daubert v. Merrill Dow Pharmaceuticals, Inc. 1993. 509 U.S. 579.

Delarosa v. Speedway, LLC. S.D. Ill. 2013. 2013 U.S. Dist. LEXIS 125104.

District of Columbia v. Heller. 2008. 554 U.S. 570.

Earp v. Novartis Pharmaceuticals Corp. E.D. N.C. 2013. 2013 U.S. Dist. LEXIS 129910.

Eastman Chemical Company v. Plastipure, Inc. 5th Cir. 2015. 775 F.3d 230.

Escola v. Coca-Cola Bottling Co. Cal. 1944. 150 P. 2d 436.

Estate of Myers v. Wal-Mart Stores, Inc. E.D. N.C. 2011. 2011 U.S. Dist. LEXIS 39164.

Exclaim Mktg, LLC v. DIRECTV, LLC. E.D. N.C. 2014 2014 U.S. Dist. LEXIS 166900.

Fernlund v. Tanscanada USA Services Inc. d/b/a Transcanada. D. Ore. 2014. 2014 U.S. Dist. LEXIS 159386.

Fisher v. United States. W.D. Wash. 2010. 2010 U.S. Dist. LEXIS 108237.

Fuentes v. Miller. N.D. Ind. 2015. 2015 U.S. Dist. LEXIS 82978.

General Electric Co. v. Joiner. 1997. 522 U.S. 136.

Hartman v. Ebsco Indus. N.D. Ind. 2013. 2013 U.S. Dist. LEXIS 140474.

Kennedy v. United States Postal Service. N.D. Ind. 2013. 2013 U.S. Dist. LEXIS 41904.

King v. Burwell. 2015. 135 S. Ct. 2480.

Kimble v. Marvel Entertainment, LLC. 2015. 135 S. Ct. 2401.

Koho v. Forest Labs, Inc. W.D. Wash 2015. 2015 U.S. Dist. LEXIS 180860.

Koho v. Forest Labs, Inc. W.D. Wash 2015. 2015 U.S. Dist. LEXIS 46050.

Kopf v. Skyrn. 4th Cir. 1993. 994 F.2d 374.

Kumho Tire Co. v. Carmichael. 1999. 526 U.S. 137.

Lemons v. Novartis Pharmaceuticals Corp. W.D. N.C. 2012. 849 F. Supp. 2d 608.

Marbury v. Madison. 1805. 5 U.S (1 Cranch) 137.

McClellan v. I-Flow Corp. D. Ore. 2010. 710 F. Supp. 2d 1092.

Mukhtar v. California State University, Hayward. 9th Cir. 2003. 299 F.3d 1053.

Olmstead v. United States. 1928. 277 U.S. 438.

Oswalt v. Resolute Industries, Inc. W.D. Wash 2012. 2012 U.S. Dist. LEXIS 190123.

Parker v. Smithfield Packing Co., Inc. E.D. N.C. 2010. 2010 U.S. Dist. LEXIS 102441.

Precision Fabrics Group, Inc. v. Tietex Int'l Ltd. M.D. N.C. 2015. 2015 U.S. Dist. LEXIS 4853.

Roberts v. Menard, Inc. N.D. Ind. 2011. 2011 U.S. Dist. LEXIS 44628.

Ruppel v. Kucanin. N.D. Ind. 2011. 2011 U.S. Dist. LEXIS 167503.

Sann v. Mastrian. S.D. Ind. 2012. 2012 U.S. Dist. LEXIS 9107.

Severn Peanut Co., Inc. v. Industrial Fumigant Co. E.D. N.C. 2014. 2014 U.S. Dist. LEXIS 34507.

SMD Software, Inc. v. Emove, Inc. E.D. N.C. 2013. 945 F. Supp. 2d 628.

Stachon v. Woodward. N.D. Ind. 2015. 2015 U.S. Dist. LEXIS 129958.

State Farm Fire & Casualty Ins. Co. v. Electrolux Home Products. N.D. Ind. 2012. 2012 U.S. Dist. LEXIS 188434.

State Farm Fire & Casualty Ins. Co. v. Electrolux Home Products. N.D. Ind. 2013. 2013 U.S. Dist. LEXIS 84550.

Suchanek v. Sturm Foods, Inc. S.D. Ill 2015. 311 F.R.D. 239.

Taylor v. Union Pacific Railroad Co. S.D. Ill 2010. 2010 U.S. Dist. LEXIS 96802.

United States v. Johnson. M.D. N.C. 2015. 122 F. Supp. 3d 272.

U.S. Automatic Sprinkler, Co., v. The Reliable Automatic Sprinkler Co. S.D. Ind. 2010. 2010 U.S. Dist. LEXIS 29456.

In re Yasmin and Yaz (Drospirenone) Marketing, Sales Practices and Products Liability Litigation. S.D. Ill. 2011. 2011 U.S. Dist. LEXIS 145552.

In re Yasmin and Yaz (Drospirenone) Marketing, Sales Practices and Products Liability Litigation. S.D. Ill. 2011. 2011 U.S. Dist. LEXIS 145593.

In re Yasmin and Yaz (Drospirenone) Marketing, Sales Practices and Products Liability Litigation. S.D. Ill. 2011. 2011 U.S. Dist. LEXIS 145989.

Winter v. Natural Resources Defense Council. 2008. 555 US 7.